

Aus dem Institut für Medizinische Statistik
der Philipps-Universität Marburg
Direktor: Professor Dr. P. Ihm

Entwicklung eines Verfahrens zur automatischen
Klassifikation von Schilddrüsenerkrankungen

Inaugural-Dissertation
zur Erlangung des Doktorgrades des
Fachbereichs Humanmedizin der
Philipps Universität Marburg/Lahn

vorgelegt von
Veronika Christiansen, geb. Schubert
aus Cottbus

Marburg 1978

Aus dem Institut für Medizinische Statistik
der Philipps-Universität Marburg
Direktor: Professor Dr. P. Ihm

Entwicklung eines Verfahrens zur automatischen
Klassifikation von Schilddrüsenerkrankungen

Inaugural-Dissertation
zur Erlangung des Doktorgrades des
Fachbereichs Humanmedizin der
Philipps Universität Marburg/Lahn

vorgelegt von
Veronika Christiansen, geb. Schubert
aus Cottbus

Marburg 1978

Angenommen vom Fachbereich Humanmedizin
der Universität Marburg am 24. 11. 1978
gedruckt mit Genehmigung des
Fachbereichs Humanmedizin

Dekan: Prof. Dr. Neurath
Referent: Prof. Dr. rer. nat. P. Ihm
Correferent: Prof. Dr. med. H. Kuni



Inhalt

1. Zur Versuchsplanung
 - 1.1 Fragestellung
 - 1.2 Untersuchungsmaterial
 - 1.3 Datenerhebung und Codierung
2. Das lernende Programm von ARKADJEV und BRAWERMAN
 - 2.1 Die Auffassung des Krankheitsbildes als "Zeichen"
 - 2.2 Die Übersetzung eines Zeichens in einen Zifferncode
 - 2.3. Patienten als Punkte im 19-dimensionalen Raum
 - 2.4 Kompakte Punktmengen im Variablenraum
 - 2.5 Die Potentialmethode
 - 2.6 Klassifikationsergebnisse bei Anwendung der Potentialmethode
 - 2.7 Diskussion der Ergebnisse
3. Taxometrische Verfahren
 - 3.1 Statistische Einführung
 - 3.1.1 Die Kovarianz als Maß für den Zusammenhang zwischen zwei Variablen
 - 3.1.2 Die "Erklärung" von Varianz
 - 3.2 Das Prinzip der Taxometrie
 - 3.2.1 Das allgemeine Faktormodell
 - 3.2.2 Das Hauptkomponentenmodell von HOTELLING
 - 3.2.3 Algebraische Grundlagen der Faktorenanalyse
 - 3.2.4 Das Hauptkomponentenmodell in der Taxometrie
 - 3.3 Die Anwendung taxometrischer Verfahren auf die Schilddrüsendiagnostik
 - 3.3.1 Ergebnisse der taxometrischen Versuche
 - 3.3.2 Die Berechnung von Diagnosewahrscheinlichkeiten
 - 3.3.3 Zur Diagnostik des autonomen Adenoms
 - 3.3.4 Vergleich des Supermerkmals z_1 mit dem klinischen Eindruck
 - 3.3.5 Diagnosewahrscheinlichkeiten und Entscheidungsregeln
4. Zusammenfassung
5. Literaturverzeichnis

1. Zur Versuchsplanung

1.1 Fragestellung

Die vorliegende Arbeit prüft verschiedene Möglichkeiten, die diagnostische Entscheidung des Arztes mit Hilfe eines Computers automatisch nachzubilden. Solche Verfahren lassen sich am ehesten für die Diagnose von Erkrankungen entwickeln und überprüfen, die sich relativ gut voneinander abgrenzen lassen und die gründlich erforscht sind. Es liegt nahe, das Gebiet der Schilddrüsenerkrankungen - im einzelnen die Krankheitsbilder Hyperthyreose, autonomes Adenom, Euthyreose und Hypothyreose - dafür auszuwählen.

Automatische Diagnoseverfahren sind vor allem dann sinnvoll, wenn schon die klinische Datenerfassung automatisch erfolgt. In der Klinik und Poliklinik für Nuklearmedizin in Marburg/Lahn wurden Verfahren entwickelt, die die Möglichkeiten der Automatisierung für die Verbesserung der Kommunikation zwischen Kliniker und einweisendem Arzt nutzen (vergl. PRIES 1972 und LÜLLWITZ, 1972).

Im Zuge dieser Entwicklung ist es ein folgerichtiger nächster Schritt, nach Möglichkeiten zu fragen, die diagnostische Entscheidungen des Arztes automatisch nachzubilden bzw. vorherzusagen. In der vorliegenden Arbeit sollen zwei Verfahren, das "lernende Programm" von ARKADJEW und BRAWERMAN und das taxometrische Verfahren hinsichtlich ihrer Leistungsfähigkeit auf dem Gebiet der automatischen Klassifikation in der medizinischen Diagnostik befragt werden.

1.2 Untersuchungsmaterial

Das Untersuchungsmaterial besteht aus Daten, die die Klinik und Poliklinik für Nuklearmedizin in Marburg/Lahn zur Verfügung stellte. Nur Patienten des Jahrgangs 1968 mit gesicherter Diagnose kamen in Frage; der Gesamtumfang der Stichprobe betrug $N = 132$ Patienten, die sich folgendermaßen auf die vier Diagnosen verteilten:

Hyperthyreose:	$N_1 = 25$
Autonomes Adenom:	$N_2 = 18$
Euthyreose:	$N_3 = 82$
Hypothyreose:	$N_4 = 7$

1.3 Datenerhebung und Codierung

Für jeden Patienten wurde folgender Ermittlungsbogen ausgefüllt:

Patientenname	Vorname
Geburtsdatum	Anschrift
Diagnose	Kliniknummer

KLINISCHE SYMPTOME

1. Gewichtsabnahme/ -zunahme in kg/Wochen
2. Fingertremor ja/nein
3. Ruhepuls Schläge/min.
4. Blutdruck mmHg
5. Hände warm/Kalt
6. Exophthalmus ja/nein
7. Hitzeempfindlichkeit ja/nein
8. T_3 -Test (Ionenaustauscher-Absorption)
9. Serumjod
- (10.) Medikamente in den letzten Wochen mit Dosisangabe
- (11.) Jod im letzten halben Jahr (z.B. Röntgenkontrastmittel)

Die Merkmale (10.) und (11.) wurden nicht codiert, aber auf besonderen Lochkarten vermerkt, da sie z.B. die Symptome 3., 4., 8., 9. beeinflussen können.

Die neun klinischen Symptome wurden in 19 Variable aufgespalten; war das Symptom nicht vorhanden, wurde es als 0 verschlüsselt, war es vorhanden, wurde es als 1 verschlüsselt.

Die neun klinischen Symptome wurden folgendermaßen aufgespalten:

1. Gewichtszunahme
2. Gewichtsabnahme
3. Puls über 100 Schläge/min.
4. Puls unter 60 Schläge/min.
5. Blutdruckamplitude größer als 60 mmHg
6. Hände warm
7. Hände kalt
8. Exophthalmus
9. Fingertremor
10. Hitzeempfindlichkeit
11. T_3 -Test unter 18 %
12. T_3 -Test zwischen 18,5 und 25 %
13. T_3 -Test zwischen 25,5 und 30 %
14. T_3 -Test über 30 %
15. Serumjod unter 3,5 $\mu\%$
16. Serumjod zwischen 3,5 und 5,4 $\mu\%$
17. Serumjod zwischen 5,5 und 7,4 $\mu\%$
18. Serumjod zwischen 7,4 und 8,0 $\mu\%$
19. Serumjod über 8,0 $\mu\%$

2. Das lernende Programm von ARKADJEW und BRAWERMAN

2.1 Die Auffassung des Krankheitsbildes als "Zeichen"

ARKADJEW und BRAWERMAN haben versucht, auf einem Computer einen Lernprozeß zur Erkennung von Zeichen nachzubilden. Unter "Zeichen" verstehen sie Mengen eines speziellen Typs: diese Mengen sind dadurch charakterisiert, daß man eine hinreichend große Anzahl von ihnen zugehörigen Elementen erkennen kann, nachdem man eine beschränkte Anzahl solcher Elemente vorher kennengelernt hat (S. 10). Die Frage ist nun, "wie ein Lernprozeß aufgebaut sein muß, damit die Maschine tatsächlich die Erkennung von Zeichen lernt. Dem Lernen soll keinerlei Mitteilung irgendwelcher Daten über diese Zeichen an die Maschine vorangehen; das Lernen selbst besteht darin, daß man der Maschine eine endliche Zahl von Objekten jedes Zeichens vorzeigt und ihr mitteilt, zu welchem Zeichen jedes Objekt gehört. Nach Abschluß des Lernens muß die Maschine fähig sein, eine hinreichend große Zahl neuer Objekte zu erkennen, die zu den gelernten Zeichen gehören." (S. 14)

Versteht man unter einem bestimmten, diagnostisch abgrenzbaren Krankheitsbild ein "Zeichen" und unter einem "Objekt" einen konkreten Patienten, der dieses Krankheitsbild aufweist, so wird das medizinische Interesse an dem von ARKADJEW und BRAWERMAN gesuchten "lernenden" Programm verständlich.

2.2 Die Übersetzung eines Zeichens in einen Zifferncode

ARKADJEW und BRAWERMAN beschäftigten sich speziell mit der Entwicklung eines lernenden Programms zur Erkennung von geometrischen Figuren; die Übersetzung eines Zeichens in einen Zifferncode soll hier jedoch von vornherein im Blick auf Symptomkonfigurationen erläutert werden.

Die erfaßten Symptome wurden alternativbinär verschlüsselt. Bei 19 erfaßten Variablen entspricht jedem Patienten dann ein Zifferncode mit 19 Stellen (ein 19-gliedriger Zeilenvektor).

Wählen wir zur Veranschaulichung nur die Variablen 8, 9 und 10 aus, so entspräche der Code (1 0 1) einem Patienten, bei dem Exophthalmus und Hitzeempfindlichkeit, nicht aber Fingertremor gefunden wurden; der Code (1 1 1) stünde für einen Patienten, bei dem alle drei Symptome ausgeprägt sind; der Code (0 0 0) für einen anderen, bei dem keines der Symptome nachgewiesen wurde; usw. Die Berücksichtigung von drei Variablen ermöglicht bei binärer Verschlüsselung theoretisch bereits $2^3 = 8$ Symptomkonfigurationen (Nämlich: (0 0 0), (1 0 0), (0 1 0), (0 0 1), (1 1 0), (1 0 1), (0 1 1), (1 1 1), vier Variable würden $2^4 = 16$ Konfigurationen, n. Variable 2^n Konfigurationen ermöglichen). (Bei den von uns erfaßten 19 Variablen ist aber nicht mit der Zahl von $2^{19} = 524288$ möglichen Konfigurationen zu rechnen, da zwischen einzelnen Variablen Abhängigkeiten bestehen: beispielsweise müßte einer 1 in der elften Variablen zwangsläufig eine 0 in der 12., 13. und 14. Variablen folgen).

Selten wird man bei zwei an derselben Krankheit leidenden Patienten identische Symptomkonfigurationen finden. Im allgemeinen wird man jedoch zwischen Symptomkonfigurationen von zwei Patienten mit derselben Diagnose eine größere Ähnlichkeit erwarten, als bei zwei Patienten mit verschiedener Diagnose. M. a. W.: die zu demselben Zeichen gehörigen Zifferncodes sollten untereinander größere Ähnlichkeit aufweisen als Zifferncodes, die zu verschiedenen Zeichen gehören. Eine wesentliche Aufgabe eines lernenden Programms besteht also darin, Ähnlichkeiten und Verschiedenheiten von Zifferncodes zuverlässig zu erkennen; der nächste Schritt bestünde dann darin, bisher unbekannte Zifferncodes dem zugehörigen Zeichen zuzuordnen, m.a.W. Patienten aufgrund ihrer Symptomkonfigurationen automatisch zu klassifizieren.

2.3 Patienten als Punkte im 19-dimensionalen Raum:

Die in 2.3 beschriebenen Zifferncodes lassen sich auch geometrisch darstellen. Wählen wir zunächst die beiden Variablen Exophthalmus x_8 und Fingertremor x_9 aus, so lassen sich die $2^2 = 4$ theoretisch möglichen Zifferncodes geometrisch als Eckpunkte eines Einheitsquadrates¹⁾ im zweidimensionalen Raum abbilden, wobei die Variablen x_8 und x_9 als Koordinaten aufzufassen sind (vgl. Abb. 1).

1) Anmerkung: Ein Einheitsquadrat hat die Seitenlänge von einer Einheit des jeweils verwendeten Maßstabs.

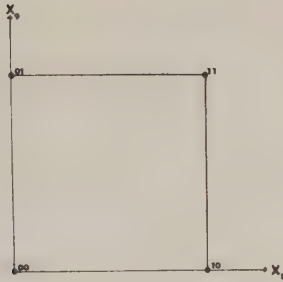


Abb. 1 Darstellung der bei zwei Variablen theoretisch
möglichen Zifferncodes als Eckpunkte eines
Einheitsquadrats im zweidimensionalen Raum.

Die gleiche Darstellungsweise ergibt bei drei
Variablen (z.B. x_8, x_9, x_{10}) die Abbildung der
theoretisch möglichen Zifferncodes als Eckpunkte
eines Einheitswürfels im dreidimensionalen Raum
(vgl. Abb. 2).

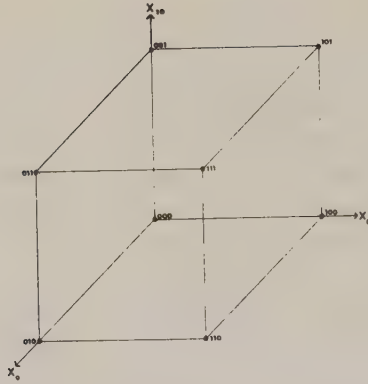


Abb. 2: Darstellung der bei drei Variablen theoretisch möglichen Zifferncodes als Eckpunkte eines Einheitswürfels im dreidimensionalen Raum.

Berücksichtigt man nun vier Variablen, so versagt die anschauliche geometrische Darstellungsweise. Dennoch spricht man in ganz analoger Weise von der Abbildung der Zifferncodes auf die Eckpunkte des vierdimensionalen "Einheitshyperwürfels". Allgemein werden also bei n Variablen die entsprechenden Zifferncodes auf die Eckpunkte des n -dimensionalen Einheitshyperwürfels abgebildet.

Würden also an jedem Patienten 19 Variable erfaßt, so entspricht jedem der aus 19 Komponenten bestehenden Patientenvektoren, also jedem aus 19 Ziffern gebildeten Code, ein eindeutig festgelegter Punkt im 19-dimensionalen Variablenraum.

2.4 Kompakte Punktmengen im Variablenraum:

Patienten mit genau identischen Zifferncodes werden auf demselben Eckpunkt des n -dimensionalen Einheitshyperwürfels gebildet. Ist jedoch mindestens eine Ziffer des Codes eines Patienten von dem eines anderen Patienten verschieden, so entspricht der in Bezug auf die betreffende Variable bestehenden Unähnlichkeiten zwischen den beiden Patienten ein Abstand im n -dimensionalen Raum.

Die Bedeutung des Abstandes soll zunächst für den Fall, daß nur drei Variable (z.B. x_8, x_9, x_{10}) erfaßt wurden, veranschaulicht werden (vgl. Abb. 2). Bei $N = 8$ Eckpunkten gibt es insgesamt $\frac{N(N-1)}{2} = \frac{8 \cdot 7}{2} = 28$ Abstände. Die Größe des Abstandes zwischen zwei Eckpunkten entspricht dem Ausmaß der Ähnlichkeit zwischen den ihnen entsprechenden Patienten. So ist z.B. ein Patient mit dem Zifferncode (1 1 1) von dem Patienten mit dem Code (0 0 0) bei Erfassung von drei Variablen maximal verschieden; dasselbe gilt für Patienten mit den Codes (1 1 0) und (0 0 1) oder (1 0 1) und (0 1 0). Wollte so ein Patient den Platz seines Gegenübers einnehmen, so müßte er jede Ziffer seines Codes mit ihrem Gegenteil vertauschen (dreifache "Inversion" des Codes), er müßte gleichsam einen Weg von drei Einheiten im dreidimensionalen Raum zurücklegen; dieser über die Anzahl der unterschiedlichen Ziffern in den Codes der entsprechenden Punkte definierte Abstand heißt in der Informationstheorie HAMMING-Abstand (vgl. Abb. 3), in der Psychologie nach F. Affneave auch City-block-Abstand; mathematisch genau ist es der Abstand in Minkowskys L_1 -Norm. (Im Einheitshyperwürfel entspricht der HAMMING-Abstand dem Quadrat des euklidischen Abstandes, was man sich mit Hilfe des Satzes von Pythagoras klarmachen kann; vgl. A & B 1966, 64 Anm. 2).

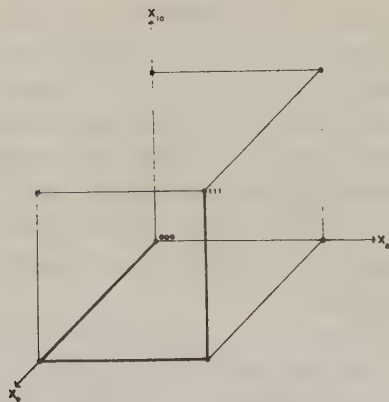


Abb. 3 Veranschaulichung des HAMMING-Abstandes zwischen den Eckpunkten (0 0 0) und (1 1 1) des dreidimensionalen Einheitswürfels.

Würde man nun beispielsweise alle Patienten mit der Symptomkostellation (0 0 0) für die Variablen x_8 , x_9 , x_{10} der Diagnose A zuteilen und alle Patienten mit dem Code (1 1 1) der Diagnose B, so betrüge der HAMMING-Abstand zwischen den beiden Diagnosen drei, der Abstand zwischen den Patienten innerhalb jeder der beiden Gruppen jedoch null Einheiten (vgl. Tab. 1, Fall 1).

Wie bereits erwähnt, ist es aber eine realistische Annahme, daß innerhalb derselben Diagnosegruppe Patienten mit verschiedenen Codes zu finden sind; rechnet man also beispielsweise auch alle Patienten mit dem Code (100) zur Diagnose A, so beträgt der Abstand zur Diagnose B jetzt nur noch mindestens zwei Einheiten, und innerhalb der Diagnose A kommen Abstände von einer Einheit vor (vgl. Tab. 1, Fall 2).

Fall	Gruppen	Codes	Mindestabstand zwischen den Gruppen	Maximaler Abstand innerhalb der Gruppe
1	A	(000)	3	0
	B	(111)		0
2	A	(000), (100)	2	1
	B	(111)		0
3	A	(000), (100)	2	1
	B	(111), (011)		1
4	A	(000), (100), (010)	1	2
	B	(111), (011)		1
5	A	(000), (100), (010)	1	2
	B	(111), (011), (101)		2

Tabelle 1: Veranschaulichung der Annäherung von zwei Diagnosegruppen bei wachsender Variabilität innerhalb der Gruppen.

Erweitert man nun Diagnose B auf alle Patienten mit dem Code (111) oder (011), so bleibt es bei dem Mindestabstand von zwei Einheiten zwischen den beiden Diagnosen, doch kommen nun auch innerhalb der Gruppe B Abstände von einer Einheit vor (vgl. Tab. 1, Fall 3).

Zählt man nun auch den Code (001) zur Diagnose A, so zeigt sich, daß diese Patienten von den zur Diagnose B gehörigen Patienten mit dem Code (011) nur um eine Einheit entfernt sind, also um weniger als von den zur eigenen Gruppe gehörigen Patienten mit dem Code (100), zu denen die Entfernung zwei Einheiten beträgt (vgl. Tab 1, Fall 4).

Fügt man schließlich noch den Code (101) der Gruppe B hinzu, so beträgt auch innerhalb der Gruppe B der maximale Abstand zwei Einheiten (vgl. Tab. 1, Fall 5; zur geometrischen Veranschaulichung vgl. Abb. 4)

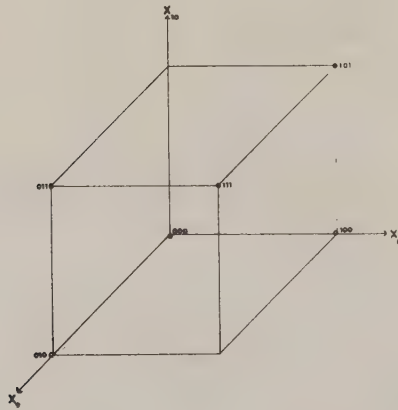


Abb. 4: Veranschaulichung der Abstände zwischen den Gruppen A und B aus Tab. 1 sowie innerhalb jeder einzelnen Gruppe.

$\circ = A, \bullet = B$

Aufgrund der Abstandskonstellation unterscheiden ARKADJEV und BRAWERMAN zwischen inneren Punkten und Randpunkten eines Zeichens. So wäre ein Punkt ein innerer Punkt einer Diagnose-Gruppe, wenn alle seine in der Entfernung Eins liegenden Nachbarpunkte entweder zu derselben oder zu gar keiner Diagnosegruppe gehören (in Abb. 4 ist also der Punkt (000) ein innerer Punkt

der Diagnosegruppe A, (111) dagegen ein innerer Punkt der Diagnosegruppe B). Beträgt jedoch der Abstand eines Punktes zu einem Punkt einer anderen Gruppe nur eine Einheit, so heißt dieser Punkt Randpunkt der Gruppe (in Abb. 4 sind die Punkte (010) und (100) Randpunkte der Gruppe A, die Punkte (011) und (101) Randpunkte der Gruppe B). Die Definitionen für innere Punkte und für Randpunkte gilt auch ganz allgemein für Punkte im n-dimensionalen Raum.

Es ist jetzt möglich, eine allgemeine Strategie für die Entwicklung eines lernenden Programms anzugeben, das in der Lage sein soll, eingegebene Symptomkonstellationen automatisch zu klassifizieren. Dem Computer sei beispielsweise die im Fall der Tabelle 1 enthaltene Information mitgeteilt worden. Wird der Computer nun so programmiert, daß er ihm unbekannte Codes demjenigen bekannten Zeichen zuordnet, zu dem sie den geringsten Abstand haben, so wird er auch die bisher unbekannten Codes (100) und (011) (Fall 2 und 3 aus Tab. 1) richtig klassifizieren. Schwierigkeiten treten jedoch bei Fall 4 und 5 auf, da das Klassifikationsprinzip für den Code (010) ebenso gut die Zuordnung zur Gruppe B erlauben würde, und der Code (101) könnte auch der Gruppe A zugeordnet werden.

Damit dürfte deutlich geworden sein, daß verschiedene Zeichen, in unserem Fall also verschiedene Diagnosegruppen umso schwerer voneinander zu trennen sind, je größer die Anzahl der in ihnen enthaltenen Randpunkte ist. M.a.W.: zwei Diagnosegruppen lassen sich umso leichter voneinander trennen, je kompakter die ihnen entsprechenden Punktmengen im Variablenraum sind; kompakt sind die Punktmengen, wenn die Anzahl

der Randpunkte im Vergleich zur Gesamtheit aller Punkte klein ist und wenn in einer hinreichend großen Umgebung fast jeden inneren Punktes nur Punkte derselben Menge liegen. Will man also mit Hilfe des genannten Klassifikationsprinzips Diagnosegruppen automatisch voneinander trennen, so geht man von der (zu überprüfenden) Hypothese aus, daß den einzelnen Diagnosegruppen kompakte Punktmengen im Variablenraum entsprechen.

2.5 Die Potentialmethode:

Das zuletzt beschriebene Klassifikationsprinzip läuft darauf hinaus, einen bisher unbekannten Punkt (bzw. Patienten) aufgrund seines Abstandes von verschiedenen bereits bekannten kompakten Punktmengen (bzw. Diagnosegruppen) der jeweils am nächsten gelegenen kompakten Punktmenge (Diagnosegruppe) zuzuordnen. Die Potentialmethode ist ein Weg, dieses Klassifikationsprinzip mathematisch zu formulieren.

Die Potentialmethode faßt einen Punkt im n-dimensionalen Variablenraum als punktförmige elektrische Ladung auf, die ein kreisförmiges elektrisches Feld um sich bildet.

Das elektrische Potential p ist dann für jeden Punkt des Raumes durch die aus der Physik bekannte Formel:

$$p = aqr^{-2} \quad \text{oder} \quad p = \frac{a q}{r^2}$$

bestimmbar, wobei a irgendein konstanter Faktor, q die Größe der jeweiligen Ladung und r der Abstand zwischen der Ladung und dem Punkt ist, für den das Potential zu bestimmen ist (Vgl. Abb. 5).

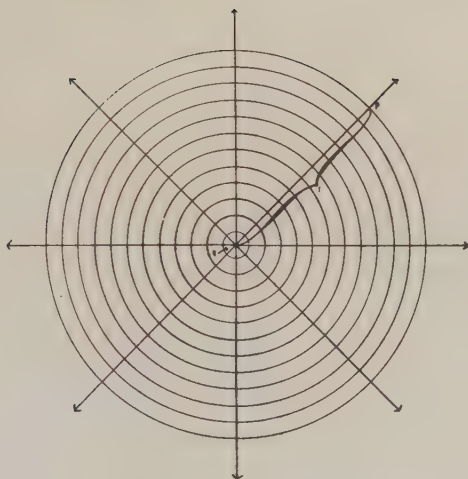


Abb. 5: Schematische Darstellung eines elektrischen Feldes.

Aus der Formel geht hervor, daß die Größe des Potentials mit steigender Entfernung von der Ladung sinkt.

Wird das elektrische Feld durch mehrere Ladungen gebildet, so ist das Potential in jedem Punkt gleich der Summe der in diesem Punkt wirksamen Einzel-potentiale. Ist also beispielsweise für einen zwischen den beiden Ladungen q_1 und q_2 befindlichen Punkt P das Potential zu bestimmen, ergibt sich:

$$p = a \frac{q_1}{r_1^2} + a \frac{q_2}{r_2^2} = a \left(\frac{q_1}{r_1^2} + \frac{q_2}{r_2^2} \right) \text{ (Vgl. Abb. 6)}$$

Wenn die Ladungen, die das Feld bilden, nun in einer kompakten Punktmenge angeordnet sind, wird das Potential im Mittelpunkt dieser kompakten Punktmenge am größten sein, und mit zunehmender Entfernung vom Mittelpunkt wird es sich verringern.

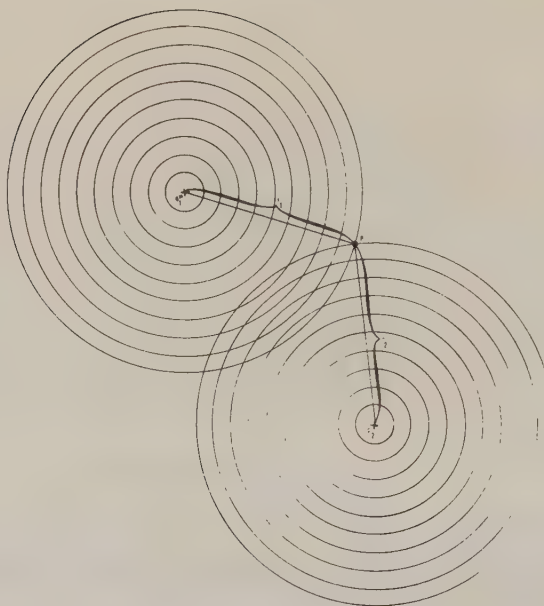


Abb. 6: Schematische Darstellung eines durch zwei Ladungen gebildeten elektrischen Feldes.

ARKADJEW und BRAWERMAN schlagen aufgrund dieser Überlegungen folgende Lösung des Klassifikationsproblems vor: "Wir ordnen jedem solchen Punkt, der im Lernprozeß auftritt, eine Funktion zu, die in ihrer Form dem elektrischen Potential analog ist, das heißt, die in diesem Punkt ihren maximalen Wert hat ²⁾ und von ihm aus nach allen Richtungen hin abnimmt (der Punkt stellt solcherart gewissermaßen die Potentialquelle dar). Eine solche Funktion kann z.B. die Funktion

$$\varphi(R) = \frac{1}{1 + \alpha R^2}$$

sein, wobei α ein Koeffizient ist, von dem das Gefälle von φ abhängt, und R die auf irgendeine

²⁾ Für $R = 0$ ist $\varphi(R) = 1$, also das Maximum der Funktion.

Weise bestimmte Entfernung zwischen dem Quellpunkt und dem Punkt ist, in dem das Potential berechnet werden soll." (S. 59) (Mit der hier vorgeschlagenen Formel ist die Schwierigkeit in der früher genannten Formel $p = a q r^{-2}$ vermieden, in der sich für $r = 0$ $p = \infty$ ergäbe.)

Die Entfernungen R eines Punktes von den Punkten einer kompakten Punktmenge lassen sich mit Hilfe des euklidischen oder HAMMING-Abstandes bestimmen. Das Potential einer kompakten Punktmenge, z.B. der Diagnosegruppe A , ist gleich der Summe der im Lernprozeß eingegebenen Einzelpotentiale. Summiert man nun die für einen fraglichen neuen Punkt geltenden Funktionswerte der einzelnen Potentiale der Diagnosegruppe A und dividiert diese Summe durch die Anzahl der Einzelpotentiale in Gruppe A (N_A), um das Resultat von der Anzahl der Quellen unabhängig zu machen, so erhält man das in dem fraglichen Punkt durch alle Quellpunkte der Gruppe A erzeugte mittlere Potential; dieses mittlere Potential charakterisiert die Entfernung des fraglichen Punktes von der Diagnosegruppe A und ist darum ein geeigneter Anhaltspunkt bei der Frage, welcher Diagnosegruppe der fragliche Punkt zuzuordnen ist.

Ist der Computer also im Lernprozeß über die Einzelpotentiale mehrerer Diagnosegruppen informiert worden, so muß er in der Erkennungsphase für jeden fraglichen Punkt die durch die einzelnen Gruppen in dem fraglichen Punkt erzeugten mittleren Potentiale berechnen; er wird den fraglichen Punkt dann der Gruppe zuordnen, deren mittleres Potential in diesem Punkt den absolut größten Wert hat.

Die Überlegungen sollen abschließend in algebraischer Form zusammengefaßt werden. Bei N_A Einzelpotentialen der Gruppe A ist die Größe des mittleren Potentials in dem zu klassifizierenden Punkt P bestimmt durch:

$$\phi_A = \frac{1}{N_A} \left(\frac{1}{1 + \alpha R_{A_1}^2} + \frac{1}{1 + \alpha R_{A_2}^2} + \dots + \frac{1}{1 + \alpha R_{A_N}^2} \right) = \frac{1}{N_A} \sum_{i=1}^{N_A} \varphi_{A_i}$$

Das durch Gruppe B erzeugte mittlere Potential ist analog dazu bestimmt durch

$$\phi_B = \frac{1}{N_B} \sum_{i=1}^{N_B} \varphi_{B_i}$$

Allgemein gilt also:

$$\phi_M = \frac{1}{N_M} \sum_{i=1}^{N_M} \varphi_{M_i} \quad , M = A, B, C, \dots$$

Die mittleren Potentiale werden nun miteinander verglichen, und der zu klassifizierende Punkt wird der Gruppe zugeordnet, deren mittleres Potential in diesem Punkt am größten ist.

Das Verfahren läuft also darauf hinaus, daß diskrete Verteilungen in stetige verwandelt werden, um so durch Interpolation auch für nicht beobachtete Werte Wahrscheinlichkeiten berechnen zu können. Durch die

Potentialfunktion wird die Wahrscheinlichkeitsverteilung interpoliert; die mittleren Potentiale (ϕ_M) werden als Wahrscheinlichkeit der jeweils in Frage stehenden Punkte angesehen.

2.6 Klassifikationsergebnisse bei Anwendung der Potentialmethode:

Die Codes der Patienten der beschriebenen Stichprobe wurden dem Computer eingegeben; gleichzeitig wurden die zu den einzelnen Codes gehörigen Diagnosen gespeichert. Um in der Klassifikationsphase mit der leichtesten Aufgabe zu beginnen, wurde dem Computer nach der Lernphase nicht ein neuer, sondern ein bereits bekannter Code eingegeben, diesmal natürlich ohne den Hinweis auf die zutreffende Diagnose. Bei diesem "Testcode" handelt es sich um eine für die Hyperthyreose sehr typische Symptomkonstellation; der Testvektor hatte die Komponenten: (0100000111000100001) (vgl. 2.3).

Der Computer war so programmiert, daß er am Ende des Rechenprogramms sofort ausdrückte, mit welchen Wahrscheinlichkeiten der Testcode den vier Diagnosegruppen zuzuordnen ist.

Die Konstante α in der Funktionsformel

$$(R) = \frac{1}{1 + \alpha R^2} \quad (\text{s.o. S. 16}),$$

die das Gefälle von σ^2 bestimmt, wurde dem Kehrwert von σ^2 gleichgesetzt ($= \frac{1}{\sigma^2}$); das Gefälle der Funktion $\varphi(R)$ wurde dann willkürlich variiert, bis eine optimale Wahrscheinlichkeitsverteilung erreicht wurde (vgl. Tab. 2). Das beste Ergebnis liefert $\sigma^2 = 0,5$; weitere Versuche, σ^2 auf einen optimalen Wert einzupendeln, erbrachten keine Verbesserung.

Diagnose	$\sigma^2 = 0,01$	$\sigma^2 = 0,10$	$\sigma^2 = 0,50$	$\sigma^2 = 1,00$
Hyperthyreose	45,44	44,66	85,68	65,82
Autonomes Adenom	20,86	21,06	12,21	17,71
Euthyreose	17,04	17,33	1,23	8,70
Hypothyreose	16,64	16,93	0,86	7,74

Tabelle 2: Wahrscheinlichkeitsverteilung in % für einen Testcode bei vier Diagnosegruppen

Da die optimale Zuordnungswahrscheinlichkeit von 85,68 % als zu niedrig anzusehen ist - zumal bei einer so typischen Symptomkonstellation -, wurde die Arbeit mit der Potentialmethode an dieser Stelle abgebrochen.

2.7 Diskussion der Ergebnisse:

Tabelle 2 zeigt, daß sowohl die Wahl eines zu großen als auch die eines zu kleinen σ^2 die Diagnosewahrscheinlichkeiten einander annähert. Läßt man σ^2 immer weiter wachsen, so geht der Wert von $\varphi = \frac{1}{1 + \frac{R^2}{\sigma^2}}$ gegen Eins; die Abstände innerhalb der kompakten Punktmengen sowie die Abstände zwischen ihnen verlieren an Bedeutung, die Potentiale fließen gewissermaßen ineinander, und kompakte Punktmengen können sich nicht mehr abheben.

Läßt man σ^2 gegen Null gehen, so wächst der Quotient $\frac{R^2}{\sigma^2}$ über alle Grenzen und geht ebenfalls gegen Null; d.h. außer in den Quellpunkten selbst, finden sich nirgends elektrische Potentiale im Raum; die Quellpunkte haben ihr maximales Potential, sind aber auch maximal isoliert, so daß sie sich nicht zu kompakten Punktmengen zusammenfügen können. Die Glättung der Kurve

bringt also einen allgemeinen Informationsverlust mit sich; ist die Streuung zu groß, fließen die Diagnosegruppen in eine einzige große Verteilung zusammen; ist sie zu klein, so gleicht die Verteilung einem Gebirge aus lauter kleinen Hügeln, zwischen denen keine Zusammengehörigkeit erkennbar ist.

Daß sich trotz optimaler Wahl von G^2 kein zufriedenstellendes Ergebnis erreichen läßt, ist wohl darauf zurückzuführen, daß die Diagnosegruppen zu viele Randpunkte im Verhältnis zur Anzahl ihrer inneren Punkte enthielten. Will man die Hypothese von der Kompaktheit der Diagnosegruppen nicht aufgeben, so ist wahrscheinlich die Art der Erhebung der Symptome für das Auftreten so vieler Randpunkte verantwortlich zu machen. Qualitative Symptome wie Tremor, Hitzeempfindlichkeit, Wärme bzw. Kälte der Hände und Exophthalmus wurden von vermutlich unterschiedlich erfahrenen, in jedem Fall aber subjektiv urteilenden Untersuchern beurteilt. Könnte man hier mit Hilfe möglichst objektiver Verfahren größere Genauigkeit erreichen, wäre es denkbar, daß die Codes trennschärfer werden.

Neben diesem grundsätzlichen Einwand ist auf den Informationsverlust hinzuweisen, den die alternativ-binäre Verschlüsselung beispielsweise bei den Symptomen Puls, Blutdruck, T_3 -Test, PBJ mit sich bringt (vgl. GROSS 1969, 127). ARKADJEW und BRAWERMAN konnten sich mit alternativer Verschlüsselung begnügen, da sie ein lernendes Programm für das Erkennen geometrischer Figuren entwickeln wollten, die sich in ausreichend feine Punktraster (nach Art eines Zeitungsbildes) zerlegen lassen. So bleibt trotz der genannten Bedenken hinsichtlich der Symptomerhebung zu untersuchen, ob eine mit kontinuierlichen Codes arbeitende Methode bessere Ergebnisse liefert.

3. Taxometrische Verfahren

Die Taxometrie gehört zu den multivariaten statistischen Verfahren; ihr Prinzip kann im Rahmen der vorliegenden Arbeit nicht vollständig dargestellt werden, weil dabei bestimmte Kenntnisse aus dem Gebiet der Regressions- und Korrelationsstatistik sowie aus der linearen Algebra vorausgesetzt werden müßten. Die folgenden Abschnitte 3.1 und 3.2 sollen lediglich ein allgemeines Verständnis für die der Taxometrie zugrundeliegende "Strategie" vermitteln, so daß es Interessierten erleichtert wird, sich anhand des dargestellten Gedankenganges und geeigneter Literatur³⁾ in die mathematischen Grundlagen der Taxometrie einzuarbeiten.

3.1 Statistische Einführung

3.1.1 Die Kovarianz als Maß für den Zusammenhang zwischen zwei Variablen

Die Frage, ob die in einer Untersuchung erhobenen Symptome wechselseitig unabhängig oder abhängig sind, gehört zu den schwierigsten Problemen der mathematisch-maschinellen Diagnostik (vgl. GROSS, S. 115). Die Methode der statistischen Erfassung von Abhängigkeiten zwischen gemessenen Symptomen soll an der Beziehung zwischen systolischem und diastolischem Blutdruck erläutert werden (Beispiel und Daten nach ÜBERLA, 1968 S. 20ff). Die an N = 10 Patienten erhobenen Meßwerte sind in Tabelle 3 zusammengestellt.

3) Korrelationsstatistik: MCCOLLOUGH & ATTA;
MITTENECKER; SACHS; WEBER. Lineare Algebra:
PAWLIK; ÜBERLA; ZURMÜHL.

Tabelle 3: Meßwerte für diastolischen und systolischen Blutdruck in mmHg bei N = 10 Patienten.

Pt.-Nr.	1	2	3	4	5	6	7	8	9	10
X-Werte (diast.)	82	72	80	72	70	82	74	78	76	78
Y-Werte (syst.)	114	116	120	110	112	130	120	125	118	114

Zeichnet man die Meßwerte in ein rechtwinkliges Koordinatensystem ein, wobei auf der X-Achse die Meßwerte für den diastolischen und auf der Y-Achse die Werte für den systolischen Blutdruck in mmHg abgetragen sind, so bilden die den Patienten entsprechenden Punkte einen annähernd elliptischen Punkteschwarm (vgl. Abb. 7). In Abb. 7 sind außerdem die arithmetischen Mittel beider Meßwertreihen ($\bar{X}$ und $\bar{Y}$) eingezeichnet:

$$\bar{X} = \frac{82+72+80+72+70+82+74+78+76+78}{10} = \frac{1}{N} \sum_{i=1}^N X_i = 76,4$$

$$\bar{Y} = \frac{114+116+...+114}{10} = \frac{1}{N} \sum_{i=1}^N Y_i = 117,9.$$

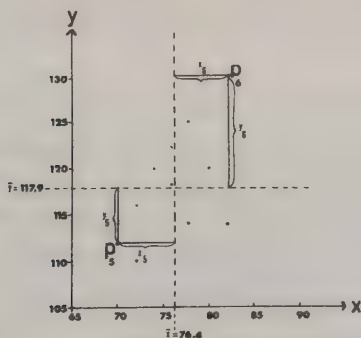


Abbildung 7: Lage der Meßwerte aus Tab. 3 im rechtwinkligen Koordinatensystem (X = diast. Blutdruck, Y = syst. Blutdruck) (nach ÜBERLA, 1968, 21)

Abb. 7 zeigt, daß bei gleichem X-Wert verschiedene Y-Werte auftreten können; es besteht also kein funktionaler Zusammenhang zwischen den beiden Variablen. (Bestünde ein funktionaler Zusammenhang, so würde jedem X-Wert genau ein Y-Wert entsprechen.)

Will man nun die Enge des Zusammenhangs zwischen zwei Variablen bestimmen, berechnet man für jeden Punkt den Abweichungswert vom Mittelwert der X-Werte und vom Mittelwert der Y-Werte, z.B. (vgl. Abb. 7)

$$x_5 = X_5 - \bar{X} = 70 - 76,4 = -6,4 \text{ oder}$$

$$y_6 = Y_6 - \bar{Y} = 130 - 117,9 = 12,1.$$

Für jeden Punkt wird nun das Produkt aus seinem Abweichungswert bezüglich X und seinem Abweichungswert bezüglich Y gebildet. Die Summe dieser Produkte wird durch die Zahl der Freiheitsgrade $N-1$ dividiert (um Vergleiche zwischen verschieden großen Stichproben zu ermöglichen). Diese auf die Zahl der Freiheitsgrade relativierte Summe der Produkte der Abweichungswerte von zwei Variablen X und Y heißt Kovarianz von X und Y:

$$\frac{(X_1 - \bar{X})(Y_1 - \bar{Y}) + (X_2 - \bar{X})(Y_2 - \bar{Y}) + \dots + (X_N - \bar{X})(Y_N - \bar{Y})}{N - 1} =$$

$$= \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{N - 1} = \frac{\sum_{i=1}^N x_i y_i}{N - 1} = s_{XY}.$$

Für das Beispiel aus Tabelle 3 ergibt sich $s_{XY} = 13,44$.

Arbeitet man mit mehr als zwei Variablen, so lassen sich entsprechend viele Kovarianzen berechnen. (Bei n Variablen gibt es $\frac{n(n-1)}{2}$ Kovarianzen). Diese Kovarianzen lassen sich in Form einer Matrix schreiben, die bei drei Variablen X, Y und Z folgendermaßen aussieht:

$$S = \begin{matrix} & \begin{matrix} (X) & (Y) & (Z) \end{matrix} \\ \begin{matrix} (X) \\ (Y) \\ (Z) \end{matrix} & \begin{pmatrix} s_{XX} & s_{XY} & s_{XZ} \\ s_{YX} & s_{YY} & s_{YZ} \\ s_{ZX} & s_{ZY} & s_{ZZ} \end{pmatrix} \end{matrix}$$

(Die obere Zeile und die linke Spalte wurden hinzugefügt, um die Symmetrie der Matrix um ihre von links oben nach rechts unten verlaufende Hauptdiagonale zu verdeutlichen.) In der Hauptdiagonalen der Kovarianzmatrix stehen gewissermaßen die Kovarianzen jeweils einer einzelnen Variablen "mit sich selbst". So ist

$$s_{XX} = \frac{\sum_{i=1}^N (X_i - \bar{X})(X_i - \bar{X})}{N - 1} = \frac{\sum_{i=1}^N X_i^2}{N - 1} - \bar{X}^2$$

die "Varianz" der Variablen X, $s_{YY} = s_Y^2$ die "Varianz" der Variablen Y usw. Die Varianz ist ein statistisches Maß für die Variabilität der Meßwerte einer Variablen um ihren Mittelwert.

In der Matrizenschreibweise steht ein einzelner Großbuchstabe (z.B. **S**) für eine gesamte Matrix von Werten. Algebraische Operationen, die sämtliche Elemente der Matrix betreffen, lassen sich mit Hilfe der Matrizensymbolik sehr übersichtlich und kurz schreiben.

3.1.2 Die "Erklärung" von Varianz

Das Ergebnis $s_{XY} = 13,44$ für die Kovarianz von diastolischem und systolischem Blutdruck gibt keinen Aufschluß darüber, ob dieser Zusammenhang als sehr eng oder eher als lose anzusehen ist. Die Enge des Zusammenhangs läßt sich jedoch beurteilen, wenn man die Abweichungswerte der beiden Variablen "standardisiert", d.h. sie in sogenannte z-Werte transformiert, indem man

sie durch ihre Standardabweichung dividiert (die Standardabweichung einer Variablen ist die Wurzel aus ihrer Varianz:

$$s_X = \sqrt{s_X^2} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2} : z = \frac{X - \bar{X}}{s_X} .$$

Indem man so die Standardabweichung der Variablen als die Einheit nimmt, in der die Abweichungen der einzelnen Meßwerte vom Mittelwert gemessen werden, macht man verschiedene Kovarianzen miteinander vergleichbar, deren einzelne Variablen ursprünglich in ganz verschiedenen Einheiten gemessen wurden. Diese standardisierte Kovarianz zwischen zwei Variablen X und Y heißt Korrelationskoeffizient von X und Y: $r_{XY} = \frac{s_{XY}}{s_X s_Y}$. Für die Meßwerte aus Tabelle 3 wurde ein Korrelationskoeffizient von $r_{XY} = 0,568$ berechnet.

Da in der Hauptdiagonalen der Kovarianzmatrix die Varianzen der einzelnen Variablen stehen, läßt sich die Kovarianzmatrix ohne weiteres in die Korrelationsmatrix überführen, indem ihre sämtlichen Elemente in der angegebenen Weise durch das Produkt aus den zugehörigen Standardabweichungen dividiert werden (in der Hauptdiagonalen stehen dann Einsen: die zu erwartenden Korrelationen der Variablen mit sich selbst). Bei standardisierten Abweichungswerten sind also Kovarianz- und Korrelationsmatrix identisch.

Der Korrelationskoeffizient ist so definiert, daß er nur zwischen -1 und +1 variieren kann: $-1 \leq r \leq +1$. Bei $r_{XY} = 0$ sind die Variablen X und Y unkorreliert, d.h. es gibt keinen (linearen) Zusammenhang zwischen ihnen: sie sind linear unabhängig voneinander.

Das Q u a d r a t des Korrelationskoeffizienten für zwei Variablen X und Y, r_{XY}^2 , heißt Bestimmtheitsmaß (auch: Determinationsindex); es bestimmt den

relativen Anteil der Gesamtvarianz, der durch den Zusammenhang zwischen zwei gemessenen Variablen bestimmt ist. (Im Beispiel aus Tabelle 3 ist die gemeinsame Varianz von diastolischem und systolischem Blutdruck $r_{XY}^2 = 0,568^2 = 0,3226$ oder 32,26 %.)

Kann man bei der Untersuchung des Zusammenhanges zwischen zwei Variablen sagen, welche Variable unabhängig (vom Untersucher manipulierbar) und welche abhängig (d.h. abhängig von der Manipulation der unabhängigen Variablen) ist, dann informiert das Quadrat des Korrelationskoeffizienten über den Anteil an der Gesamtvarianz, der durch die unabhängige Variable "erklärt" ist. Untersucht man beispielsweise die Abhängigkeit des Bremsweges von der Geschwindigkeit, so wären bei einer gefundenen Korrelation von $r = 0,8$ 64 % der Varianz der in der Untersuchung gemessenen Bremswege durch den Einfluß der Geschwindigkeit erklärt. Je höher also die gefundene Korrelation ist, desto genauer wird man für bestimmte Werte der unabhängigen Variablen die zu erwartenden Meßwerte der abhängigen Variablen vorhersagen können. Man kann also sagen, daß die Projektionen der Meßwertpunkte auf die X-Achse (auf die Koordinate "Geschwindigkeit") ihre Lage in Bezug auf die Y-Achse (den Bremsweg) bereits zu einem beträchtlichen Teil festlegt.

Die Kovarianzmatrix (bzw. die Korrelationsmatrix) ist der Ausgangspunkt des taxometrischen Rechteinverfahrens; die vorstehenden Hinweise sollten verdeutlichen, welche Informationen in ihr enthalten sind. Berechnet man die Elemente der Kovarianzmatrix aus den in der jeweiligen Stichprobe erhobenen Meßwerten, so hat man Informationen über die Varianz (s_x^2) und Standardabweichung (s_x) jeder

einzelnen Variablen, über das Ausmaß des Zusammenhanges jeder Variablen mit jeder anderen (s_{XY} bzw. r_{XY}) und über den relativen Anteil der Gesamtvarianz der durch den zwischen je zwei Variablen gefundenen Zusammenhang bestimmt ist (r_{XY}^2).

3.2 Das Prinzip der Taxometrie

Man kann die Taxometrie in bestimmter Hinsicht als eine Umkehrung faktorenanalytischer Verfahren auffassen. Das Konzept der Faktorenanalyse soll darum in den Grundzügen kurz referiert werden, um die spezifische Abwandlung des Modells in der Taxometrie besser verständlich werden zu lassen.⁴⁾

3.2.1 Das allgemeine Faktormodell

"Unter F a k t o r e n a n a l y s e versteht man ein mehrdimensionales Modell, das die Analyse quantitativer (ausgezählter, gemessener) Merkmale in funktional einfachere Grundvariable oder F a k t o r e n gestattet." (PAWLIK, S. 17) Das Verfahren ermöglicht es, Zahl und Art jener "Elementarvariablen" oder "Faktoren" zu bestimmen, als deren Funktion die untersuchten Merkmale darstellbar sind. "Da die Zahl dieser Faktoren stets erheblich kleiner ist als die Zahl der Beobachtungsmerkmale, von denen ausgegangen wurde, bietet sich in den Faktoren ein ungleich sparsameres und handlicheres Bezugssystem zur Darstellung eines Merkmalsbereiches an." (PAWLIK, S. 18)

Innerhalb der Faktorenanalyse ist zwischen dem allgemeinen Faktormodell und den daraus ableitbaren speziellen Faktormodellen (von denen eins im Abschnitt 3.2.2 behandelt wird) zu unterscheiden:

4)Referat nach PAWLIK, 1968; Literatur bei PAWLIK, 1968 und ÜBERLA, 1968. Voraussetzung: lineare Algebra; Einführung bei PAWLIK und ÜBERLA.

"Das allgemeine Faktormodell definiert in wenigen Grundgleichungen das Modell der Aufspaltung der Varianz eines quantitativen Merkmals in nicht direkt beobachtbare Elementarvariable. Um die Eindeutigkeit der Lösung sicherzustellen, müssen darüber hinaus zusätzliche Annahmen eingeführt werden. Je nach der Wahl dieser Zusatzannahmen, erhält man ein anderes spezielles Faktormodell." (PAWLIK, S. 18)

Die Faktorenanalyse faßt jedes quantitative Merkmal X als eine lineare Funktion mehrerer Faktoren auf. Die "Bestimmungsgleichung" des Merkmals X in den r Faktoren definiert X als eine gewichtete Summe von Faktoren:

$$X = a_1 f_1 + a_2 f_2 + \dots + a_r f_r,$$

wobei der Buchstabe f für die r Faktoren und a für die den jeweiligen Faktoren zugehörigen Gewichts-
zahlen steht. Die Faktoren sollen "Elementarvariablen" und darum wechselseitig statistisch unabhängig sein, d.h. ihre Interkorrelationen sollen sämtlich Null sein:

$$r_{f_p f_q} = 0 \text{ für } p \neq q = 1, 2, \dots, r.$$

In der genannten Bestimmungsgleichung des Merkmals X sind individuelle Differenzen zwischen den Personen in den Merkmals u n spezifischen "Faktorwerten", d.h. in den Meßzahlen der Personen in den Faktoren $f_1, f_2, \dots, f_r$ vorgesehen. Dagegen sind die Gewichts-
zahlen a in der Bestimmungsgleichung die merkmals-
spezifischen "Faktorenladungen" des Merkmals X . Ein Faktor f_p ist ohne Einfluß auf das Merkmal X , wenn die Faktorenladung a_p von X um Null liegt, und er hat umso größeren Einfluß, je höher X im Faktor f_p "geladen" ist.

"Das Ergebnis einer Faktorenanalyse wird Faktorenlösung genannt. Ihr sind die Faktorenladungen der Variablen und die Faktorenwerte der $V_{pn}^{5)}$ zu entnehmen. Man nennt eine solche Lösung vollständige Faktorenlösung der m Variablen, wenn aus ihr die gesamte Varianz jeder Variablen erklärt wird (die experimentellen Bedingungen exakt reproduzierbar sind)." (PAWLIK, 1968, 36).

Muß dazu die Faktorenzahl im allgemeinen mindestens gleich der Zahl der Variablen sein, so bleiben in der reduzierten Faktorenlösung diejenigen Faktoren bei der Ergebnisinterpretation unberücksichtigt, die nur einen geringen Anteil der Gesamtvarianz "erklären". "Zum Unterschied von der vollständigen gestattet die reduzierte Faktorenlösung freilich nur eine näherungsweise Bestimmung der Versuchsergebnisse aus den Faktoren." (PAWLIK, 1968, 36).

3.2.2 Das Hauptkomponentenmodell von HOTELLING

Eine Faktorenanalyse nach dem Hauptkomponentenmodell läßt sich auf eine Rotation des Koordinatensystems im Variablenraum zurückführen. Sind die $r = m$ Faktoren der vollständigen Lösung nach fallendem Varianzbeitrag gereiht, so daß $V_p \geq V_{p+1}$, so besteht die reduzierte Faktorenlösung dann einfach aus den ersten k der r Faktoren. "Der reduzierte Faktorenräum einer Analyse nach dem Hauptkomponentenmodell ist ein k -dimensionaler Unterraum des vollständigen Faktorenraumes und damit auch des Variablenraumes." (PAWLIK, 1968, 60).

5) V_{pn} = Abkürzung für "Versuchspersonen".

Statt der Variablen bestimmt eine Faktorenanalyse nach dem Hauptkomponentenmodell also ein neues rechtwinkliges Koordinatensystem, dessen eindeutige Lage durch die Forderung gesichert wird, daß die k Faktoren ($1 \leq k \leq m$) ein Maximum der totalen Gesamtvarianz ausschöpfen: "Das Verhältnis des Varianzbeitrages eines Faktors zur noch nicht erklärten totalen Merkmalsvarianz soll ein Maximum sein. ... Relativ zu anderen Systemen von k unkorrelierten Variablen geben die so bestimmten Faktoren die erschöpfendste Erklärung der Beobachtungsmerkmale." (PAWLIK, 1968, 61).

Die Matrizengleichung für die Hauptachsenrotation läßt sich auf m Variablen verallgemeinern; bei drei Variablen bilden die Personenpunkte ein dreidimensionales Ellipsoid, bei m Variablen ein m -dimensionales "Hyperellipsoid". "So wie die Ellipse zwei, das Ellipsoid drei, besitzt das Hyperellipsoid m verschieden lange Hauptachsen. In der Hauptachsenmethode wird nun der erste Faktor kollinear mit der längsten dieser m Hauptachsen gelegt, f_2 kollinear mit der zweitlängsten usf. bis f_m , der mit der kürzesten Hauptachse kollinear ist." (PAWLIK, 1968, 63).

3.2.3 Algebraische Grundlagen der Faktorenanalyse

Die algebraischen Grundlagen der Faktorenanalyse können hier nur kurz gestreift werden (Einzelheiten bei PAWLIK, 1968 und ÜBERLA, 1968). Ausgangspunkt ist das Dispersionsellipsoid standardisierter Variabler. Gesucht wird die Hyperebene, die alle bzw. fast alle Informationen enthält. Diese Hyperebene wird durch die von Null verschiedenen Hauptachsen des Dispersionsellipsoides bestimmt.

Ist Z die Datenmatrix (k -te Spalte gleich Vektor des j -ten Individuums) und sind die Variablen streuungsnormiert mit Mittelwert Null, so ist die Korrelationsmatrix

$$R = Z Z'.$$

Z' ist die transponierte Matrix. Die Faktorenanalyse faßt jedes quantitative Merkmal als eine gewichtete Summe mehrerer Faktoren auf; das läßt sich schreiben als

$$Z = A F,$$

wobei A die Faktorenmatrix heißt und die Faktorenladungen enthält; F ist die Faktorwertematrix.

Einsetzen für Z ergibt

$$Z Z' = (A F) (A F)' = A A'.$$

A hat orthogonale Spalten:

$$A' A = \Lambda.$$

$\Lambda = (\lambda_i)$ ist die Diagonalmatrix der Eigenwerte. Bezeichnet man die Spalten von A mit a_j , so ist a_j, λ_j Lösung der Gleichung

$$R a = \lambda a,$$

$$|a| = \sqrt{\lambda}$$

(Eigenwertsproblem).

Ist R von der Ordnung n , gibt es ein Lösungssystem von n Werten

$$\lambda = \lambda_1, \lambda_2, \dots, \lambda_n$$

$$a = a_1, a_2, \dots, a_n.$$

Die a_j geben die Hauptachsen des durch R dargestellten Dispersionsellipsoides wieder und stellen ihrerseits ein Koordinatensystem dar, das dem ursprünglichen gegenüber gedreht ist. Bezüglich dieses Systems läßt sich jedes Individuum nun durch Koordinaten p darstellen, wobei

$$Z = A' P$$

gilt. Die k -te Spalte von P stellt die Faktorenwerte für das k -te Individuum dar. Die Komponenten von P sind unkorreliert und haben Einheitsvarianz.

3.2.4 Das Hauptkomponentenmodell in der Taxometrie

Während man in der Faktorenanalyse voraussetzt, daß die Stichprobe der n untersuchten Personen homogen ist, geht man in der Taxometrie von der gegenteiligen Annahme, also von der Inhomogenität der Stichprobe aus. Es kommt dann darauf an, die Gesamtstichprobe in Teilstichproben zu zerlegen, die den vermuteten einzelnen Gruppen entsprechen; man sucht also Gruppen auf, die im Voraus nicht bekannt sind. Die Taxometrie bedient sich dabei der faktorenanalytischen Technik, doch richtet sich ihr Hauptinteresse nicht so sehr auf das Auffinden der varianzstärksten Faktoren zur Erklärung der Gesamtvarianz, sondern auf die Trennung von Gruppen anhand der Projektionen der Personenpunkte auf diese varianzstärksten Faktoren.

Man geht im allgemeinen nicht von der Korrelations-, sondern von der Kovarianzmatrix aus, doch ist die Technik im übrigen der der Hauptkomponentenmethode analog.

Beim einfachsten taxometrischen Modell setzt man voraus, daß innerhalb der Gruppen die Variablen unkorreliert sind und die gleiche Varianz haben. Bei der Hauptkomponentenanalyse wird analog verfahren, die Varianz der Faktorenwerte wird lediglich auf Eins normiert.

In der Taxometrie werden die Individuen durch die Faktorenwerte bezüglich der Hauptkomponenten abgebildet, jedoch mit der Abänderung, daß die Varianz nicht auf Eins, sondern auf λ_i normiert wird, d.h. man verwendet die Faktorenwerte multipliziert mit $\sqrt{\lambda_i}$.

Im einfachsten Fall stellt man sich vor, daß die Personenpunkte pro Gruppe im m -dimensionalen Raum

eine hypersphärische Punktwolke bilden. "Man sucht durch die Gruppenschwerpunkte eine Ebene zu legen und projiziert die Punkte senkrecht auf diese Ebene, um dort die Gruppen zu suchen." (IHM, TRAUTNER und WOLF, 1971). Die Personenpunkte werden also den Forderungen des Hauptkomponentenmodells entsprechend optimal auf die Ebene der größten Varianz projiziert, die durch die k längsten Halbachsen des m -dimensionalen Dispersionshyperellipsoides bestimmt ist; sind g Untergruppen vorhanden, wäre eine $(g - 1)$ -dimensionale Ebene zur Darstellung ohne Informationsverlust ausreichend; bei $g \geq 4$ spricht man von Hyperebenen. Aus praktischen Gründen projiziert man jedoch nacheinander auf die durch je zwei der drei längsten Halbachsen bestimmten Ebenen (vgl. IHM et al., 1967; IHM und LIEBAU, 1965; IHM, TRAUTNER und WOLF, 1971).

3.3 Die Anwendung taxometrischer Verfahren auf die Schilddrüsendiagnostik

3.3.1 Ergebnisse der taxometrischen Untersuchung

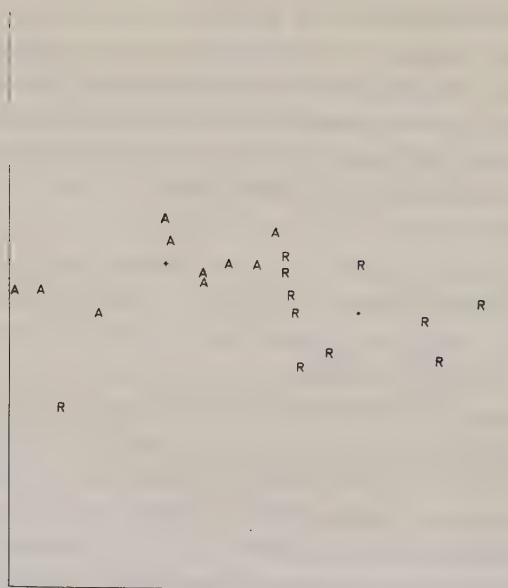
Das taxometrische Verfahren wurde in seinen Grundzügen dargestellt. Die Patienten werden durch Punkte in einem m -dimensionalen Raum dargestellt, und derart in einen dreidimensionalen Raum projiziert, daß die erste Koordinatenachse der Richtung der größten Varianz entspricht, die zweite der Richtung der nächstgrößten Varianz senkrecht zu der ersten Achse usw.

Die vorliegenden Daten haben eine m -dimensionale stetigdiskrete Mischverteilung, für die es wegen der Abhängigkeit der einzelnen Merkmalsgrößen keine befriedigende Modellvorstellung gibt. Unter der Annahme gleicher Kovarianzmatrizen bei den drei Gruppen (Diagnosen) kann man jedoch eine lineare Transformation der Merk-

malsgrößen in der Weise vornehmen, daß die Binnen-gruppenkovarianzmatrix zu einer Einheitsmatrix wird ("Mahalanobisation"). Die Projektion auf eine Ebene, die durch die (transformierten) Gruppenschwerpunkte geht, enthält dann alle Information über die einzelnen Gruppen.

Zunächst wurden alle Daten aus der erhobenen Stichprobe taxometrisch verarbeitet; die taxometrische Projektion zeigte eine unübersichtliche Vermischung der vier Diagnosegruppen Hyperthyreose (Hy), autonomes Adenom (Ad), Euthyreose (Eu) und Hypothyreose; lediglich 14 Hyperthyreosen hoben sich von der Gesamtgruppe ab.

Der zweite Trennversuch beschränkte sich auf die zwei Diagnosegruppen Hy und Ad. Um möglichst homogene Gruppen zu erhalten, wurden nur weibliche Patienten berücksichtigt. Außerdem wurden jetzt die "harten" Daten Ruhepuls, Blutdruck, T_3 -Test und PBJ-Wert nicht mehr alternativ-binär, sondern quantitativ codiert. Nur die Gewichtsveränderung in den letzten zwei Monaten wurde weiterhin alternativ verschlüsselt (Abnahme: -1, unverändert: 0, Zunahme: 1). Der Ruhepuls dagegen wurde in Schlägen pro Minute angegeben. Für systolischen und diastolischen Blutdruck wurden die Abweichungen vom Mittelwert der entsprechenden Alters- und Geschlechtsgruppen ermittelt; als Maßeinheit diente die Standardabweichung; Mittelwert und Standardabweichung zur Korrektur der Blutdruckwerte wurden den GEIGY-Tabellen entnommen. Auf diese Weise wird der lineare Zusammenhang zwischen Alter und Geschlecht einerseits und Blutdruck andererseits berücksichtigt. Das Ergebnis im T_3 -Test wurde in Prozent Ionenaustauscherabsorption gemessen, der PBJ-Wert wurde mit



HORIZONTAL ENTSpricht ERSTEM EIGENVEKTOR

VERTIKALE ENTSpricht ZWEITEM EIGENVEKTOR

Abbildung 8: Projektion von Hy (R) (N = 13) und Ad (A) (N = 10) für weibliche Patienten. Gute Trennung von Hy einerseits und Ad andererseits. R (Hy) am äußersten Ende von A (Ad) ist ein Ausreißer, der sich als Fehldiagnose herausstellt.

dem Signette-Verfahren bestimmt, das Serumjod mit dem Autoanalyzer).

Das Ergebnis des zweiten Versuchs zeigt eine gute Trennung der beiden Diagnosegruppen sowie einen deutlichen Hyperthyreose-Ausreißer am äußersten Extrem der Gruppe Ad; dieser Ausreißer erwies sich bei Kontrolle als Eu (typisches Beispiel für die Ausreißerkontrolle von IHM und LIEBAU, 1965); das Ergebnis des zweiten Versuchs ist in Abbildung 8 dargestellt.

Im dritten Versuch wurden die Daten der weiblichen Euthyreose-Patienten mit einbezogen; zwei extreme Fälle von Eu blieben unberücksichtigt. Abbildung 9 zeigt, daß zwei Fälle von Eu zu weit unterhalb der Gruppen Eu und Ad liegen; die Kontrolle ergab, daß es sich um Hypertonien handelt.

Abbildung 10 zeigt die drei Diagnosegruppen ohne die beiden Hypertonien aus dem dritten Versuch: Hy einerseits und Ad und Eu andererseits sind gut getrennt.

Führt man das Projektionsverfahren nun für die Diagnosen Ad und Eu durch, zeigt sich, daß die beiden Gruppen aufgrund der verwendeten Merkmale nicht getrennt werden können (vgl. Abbildung 11).

Die taxometrische Versuchsserie hat also ergeben, daß die ursprünglich vier Diagnosegruppen aufgrund der erhobenen Daten nicht zu trennen sind. Doch läßt sich die Gruppe Hy einerseits von Ad und andererseits von Eu trennen; Ad und Eu dagegen lassen sich kaum trennen.

Auf die Trennung zwischen Hy einerseits und Ad und Eu andererseits ist im folgenden noch einzugehen. An dieser Stelle ist festzuhalten, daß durch die taxometrische Projektion extreme Hyperthyreosen (Abbildung 8)

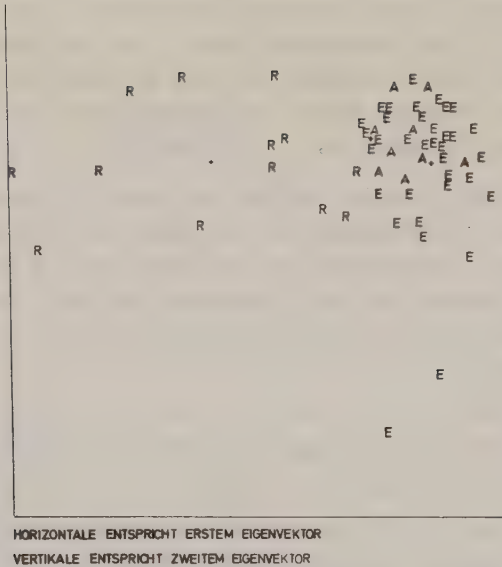


Abbildung 9: Projektion von Hy (R) (N = 13), Ad (A) (N = 10) sowie von Eu (N = 39) für weibliche Patienten
Gute Trennung von Hy (R) einerseits zu Ad (A) und Eu (E) andererseits. Allerdings wurden zwei extreme Fälle von Eu nicht mit berücksichtigt.
Zwei Ausreißer (E) unterhalb von Eu (E); es handelt sich um fehldiagnostizierte Hypertonien.

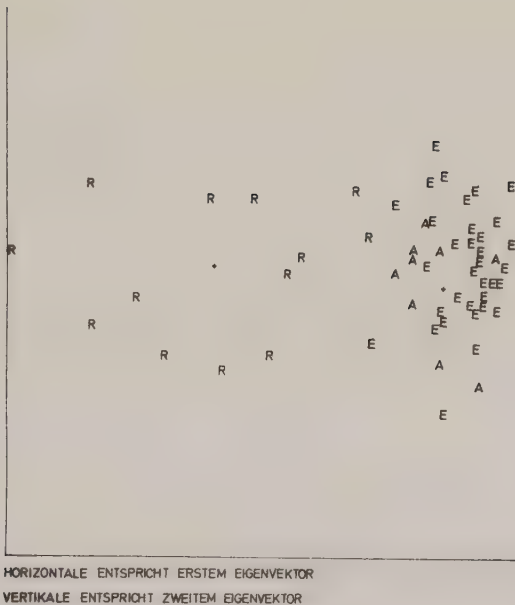


Abbildung 10: Projektion von Hy (R) (N = 13), Ad (A) (N = 10) und Eu (E) (N = 37) für weibliche Patienten ohne die beiden fehldiagnostizierten Hypertonien. Gute Trennung von Hy (R) einerseits und Ad (A) und Eu (E) andererseits.

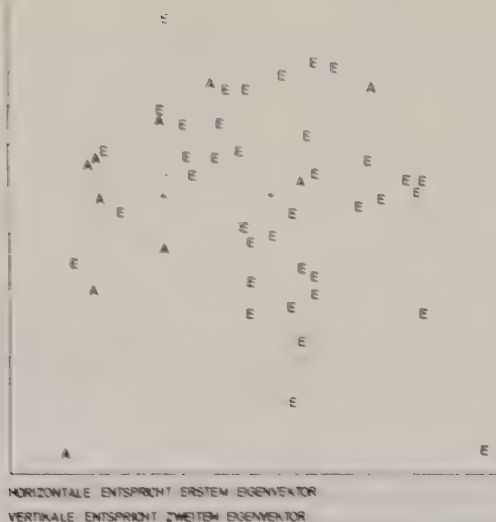


Abbildung 11: Projektion von Ad (A) ($N = 10$) und Eu (E) ($N = 41$) für weibliche Patienten unter Miteinbeziehung der fehldiagnostizierten Patienten; hier ergab sich keine ausreichende Trennung der beiden Diagnosegruppen aufgrund der verwendeten Merkmale.

und Euthyreosen (Abbildung 9) erkannt werden; Fehldiagnosen projizieren sich als Ausreißer und können mit Hilfe des Verfahrens ausgeschieden werden.

Inzwischen hat PROKOPH ebenfalls bei Untersuchungen zur Schilddrüsendiagnostik an 184 Patienten mehr als 10 Fehldiagnosen mit Hilfe des taxometrischen Verfahrens aufdecken können (PROKOPH, 1973, 190.193).

PROKOPH fand außerdem bei seinen Untersuchungen durch taxometrische Projektion eine neue Diagnosegruppe ("Struma stigmatica"); bei Patienten dieser Gruppe zeigt die Schilddrüse keinen auffälligen Tastbefund wie bei einer Struma diffusa, doch finden sich im Szintigramm auffällige Speicherbezirke wie bei einer Struma nodosa. Die Abtrennung dieser Diagnosegruppe basiert allein auf den erfaßten Symptomen; die Frage der Pathogenese bleibt davon unberührt (PROKOPH, 1973, 194). PROKOPH fand weiterhin, daß Frauen mit vielen Geburten anfälliger für Schilddrüsenerkrankungen sind als Frauen mit wenigen Geburten (PROKOPH, 1973, 194).

Das taxometrische Verfahren ist also in der Lage, Ausreißer bzw. Fehldiagnosen aufzudecken und neue nosologische Einheiten zu finden.

3.3.2 Die Berechnung von Diagnosewahrscheinlichkeiten

Es hat sich gezeigt, daß die Patienten der drei Diagnosegruppen Hy, Ad und Eu sich in einem zweidimensionalen Koordinatensystem darstellen lassen, dessen Achsen den HOTELLING'schen Hauptachsen entsprechen. Im Sinne der linearen Diskriminanzanalyse FISHER's genügen zwei Supermerkmale zur Darstellung der Diagnosegruppen. Die Koordinaten der Projektion z_1 und z_2

entsprechen diesen Supermerkmalen, die ihrerseits Linearkombinationen der ursprünglichen Merkmalsgrößen sind. Wenn diese Merkmalsgrößen auch nicht unabhängig voneinander sind, so sind z_1 und z_2 unter recht allgemeinen Voraussetzungen doch annähernd normal verteilt, so daß die für die Normalverteilung gültigen, auf dem BAYES'schen Theorem beruhenden statistischen Entscheidungsverfahren anwendbar sind. Setzt man als Verteilung von z_1 und z_2 Normalverteilung voraus, ist zu berücksichtigen, daß deren Parameter nur aus Stichproben relativ kleinen Umfangs bekannt sind. Darum führt der BAYES'sche Ansatz auf eine zweidimensionale Student-Verteilung, d.h. auf eine Verallgemeinerung der t-Verteilung (vgl. IHM et al., 1970, 12ff).

Bezeichnet H die Diagnose und N den Umfang einer Stichprobe von Patienten mit Diagnose H, so ist die Wahrscheinlichkeit von H

$$L(H) = \frac{N(N-2)}{2\pi(N-1)(N+2)} \frac{1}{s_1 s_2 \sqrt{1-r^2}} \left(1 + \frac{1}{r} \frac{N}{N+1} Q^2\right)$$

wobei s_1^2 die Varianz von z_1 , s_2^2 die von z_2 und r der Korrelationskoeffizient ist. Für Q^2 gilt

$$Q^2 = \frac{1}{1-r^2} \left\{ \frac{(z_1 - \bar{z}_1)^2}{s_1^2} - \frac{2r(z_1 - \bar{z}_1)(z_2 - \bar{z}_2)}{s_1 s_2} + \frac{(z_2 - \bar{z}_2)^2}{s_2^2} \right\}$$

Hierbei sind z_1, z_2 die beim Patienten beobachteten Werte, $\bar{z}_1, \bar{z}_2$ die vorher bei der Stichprobe von N Patienten bekannter Diagnose ermittelten mittleren z_1 und z_2 .

Schreibt man die drei Diagnosen Hy, Ad und Eu als H_1 , H_2 und H_3 , so lassen sich die Wahrscheinlichkeiten der jeweiligen Diagnose als $P(H_i)$ für $i = 1, 2, 3$ nach der Formel

$$P(H_i) = \frac{L(H_i)}{L(H_1) + L(H_2) + L(H_3)}$$

berechnen. Ist beispielsweise $P(H_1)$ sehr groß,

$$P(H_1) = 1 - \alpha,$$

so kann man die Hypothese, daß Hy vorliegt, mit subjektiver Wahrscheinlichkeit von $1 - \alpha$ annehmen.

Aus den Verteilungen von z_1 und z_2 kann man für die einzelnen Diagnosegruppen je eine Ellipse so berechnen, daß mit subjektiver Wahrscheinlichkeit von $1 - \alpha$ der Punkt (z_1, z_2) im Inneren dieser Ellipse liegt (Einzelheiten bei IHM et al., 1970, 12ff).

In Abbildung 12 sind diese Ellipsen für $\alpha = 0.01$ eingezeichnet. Außerdem sind die Kurven eingetragen, auf deren Punkten die Wahrscheinlichkeit für eine Hyperthyreose 99 % bzw. 95 % beträgt. Berechnet man nun für einen Fall unbekannter Diagnose die Koordinaten z_1 und z_2 und zeichnet den entsprechenden Punkt ein, so ist für alle Punkte auf oder rechts von der Kurve die Wahrscheinlichkeit für eine Hyperthyreose 99 %. Liegt ein Punkt außerhalb aller Ellipsen, repräsentiert er vermutlich einen Ausreißer, für den das Modell nicht gilt.

Abbildung 12 zeigt, daß die Ellipse für Hy praktisch die für Ad und Eu einschließt. Die geringe Information aus der kleinen Stichprobe bringt es mit sich, daß die Diagnose Hy zwar in vielen Fällen

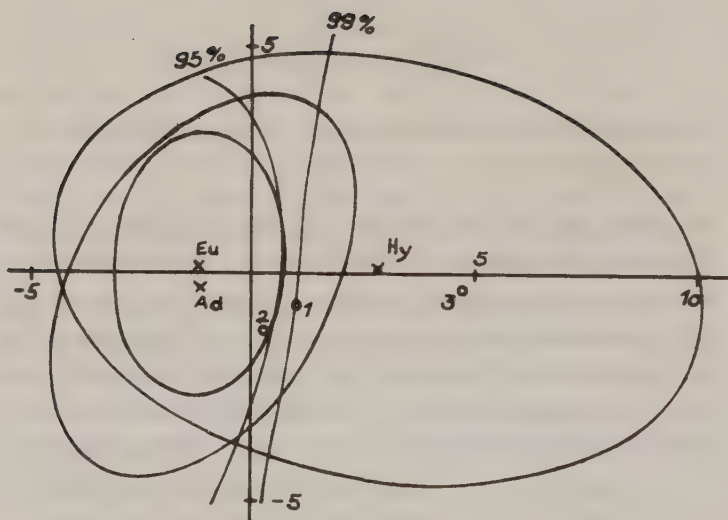


Abb. 12: Koordinatensystem mit den Supermerkmalen z_1 und z_2 als Achsen. Eingezeichnet wurden die Gruppenschwerpunkte für Hy, Ad und Eu, die Streuungsellipsen, innerhalb derer 99 % der entsprechenden Fälle liegen, sowie die Grenzen für die Wahrscheinlichkeit von Hy bei 95 % und 99 % Sicherheit.

mit 95- bzw. 99prozentiger Sicherheit behauptet werden kann, doch läßt sie sich im Vergleich zu Ad und Eu kaum ausschließen.

Das Verfahren wurde an einer Reihe unausgelesener Fälle erprobt, bei denen die Diagnose gesichert war; diese Erprobungsstichprobe umfaßte 14 Hyperthyreosen, 4 autonome Adenome und 13 euthyreote vegetative Dystonien. Die für diese Fälle errechneten Diagnosewahrscheinlichkeiten sind in Tabelle 4 wiedergegeben.

Tabelle 4: Diagnosewahrscheinlichkeiten (in %) bei
31 diagnostisch gesicherten Schilddrüsenerkrankungen (Erprobungsstichprobe):

gesicherte Diagnose	errechnete Diagnosen		
	Hy	Ad	Eu
Eu	21	28	61
Eu	39	29	32
Eu	32	23	45
Eu	11	20	69
Eu	25	6	69
Eu	18	18	64
Eu	14	16	70
Eu	17	34	48
Eu	26	15	59
Eu	66	11	23
Eu	13	30	57
Eu	10	21	69
Eu	78	3	19
Hy	99	1	-
Hy	100	-	-
Hy	100	-	-
Hy	100	-	-
Hy	100	-	-
Hy	100	-	-

Hy	100	-	-
Hy	100	-	-
Hy	100	-	-
Hy	100	-	-
Hy	100	-	-
Hy	100	-	-
Hy	100	-	-
Hy	100	-	-
Ad	79	6	15
Ad	47	7	46
Ad	38	5	57
Ad	14	23	63

BAYES'sche Verfahren lassen sich aufgrund der Erfahrung verbessern. Aus diesem Grunde wurden die 34 gesicherten Diagnosen der Erprobungsstichprobe in die Berechnungen einbezogen; zusammen mit der Ausgangsstichprobe ergaben sich so folgende Diagnosehäufigkeiten: Hy = 27, Ad = 14, Eu = 52. Abbildung 13 zeigt die Ellipsen für die neuen Wahrscheinlichkeitsverteilungen sowie die 99 %-Kurve für die Diagnose Hy.

Abbildung 14 zeigt die Ellipsen sowie die 95- und 99 %-Kurven, die sich ergeben, wenn für die Berechnung der Wahrscheinlichkeiten nur die sieben klinischen Merkmale herangezogen werden. Vergleicht man Abbildung 14 mit Abbildung 13, also die Ellipsen für sieben bzw. neun Merkmale, so sieht man deutlich die verbesserte Entscheidungsmöglichkeit, wenn T_3 -Test und PBJ-Wert in die Untersuchung mit einbezogen werden. Mißt man die Fläche der Hy-Ellipse rechts neben der 99 %-Kurve, ergeben sich bei Verwendung der sieben klinischen Merkmale 41% der Gesamtfläche, bei neun Variablen jedoch 69%. Von allen Hyperthyreosen werden mittels der klinischen u n d nuklearmedizinischen

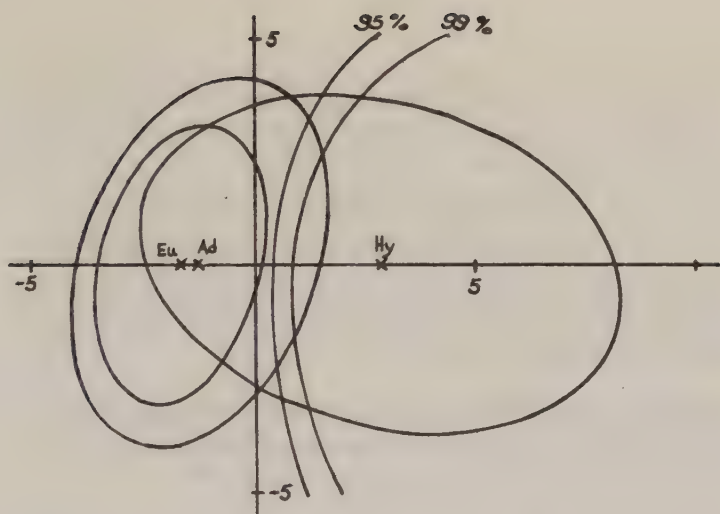


Abb. 13: wie Abb. 12, jedoch zweite Berechnung mit zusätzlichen bestätigten Diagnosen (Lernprozeß)

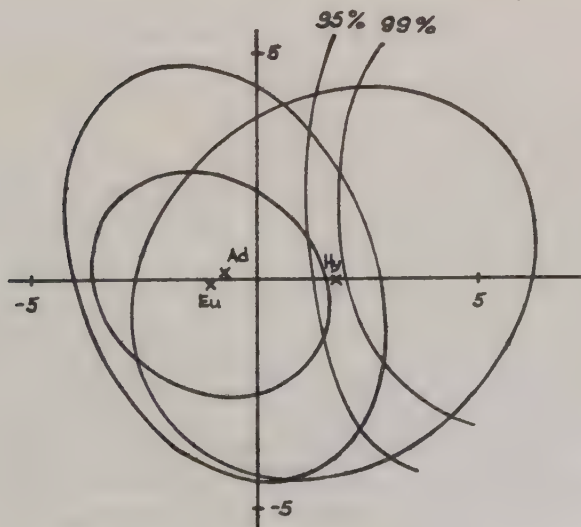


Abb. 14: wie Abb. 13, jedoch nur für die sieben klinischen Merkmale

Merkmale also ca. 28 % mehr Fälle auf dem 99 %-Niveau entschieden als mit den klinischen Merkmalen allein.

3.3.3 Zur Diagnostik des autonomen Adenóms

Bei der Erprobung des taxometrischen Verfahrens wurden zusätzlich die Diagnosewahrscheinlichkeiten für drei Fälle von dekompensiertem autonomen Adenom berechnet (Tabelle 5).

Tabelle 5: Diagnosewahrscheinlichkeiten für drei dekompensierte autonome Adenome (in %):

Fall	Hy	Ad	Eu
1	99	1	-
2	92	3	5
3	100	-	-

Bei dekompensierten autonomen Adenomen ist im Gegensatz zu den kompensierten euthyreoten das kybernetische Funktionssystem gestört oder bereits zusammengebrochen. Die große Wahrscheinlichkeit für Hy in den Fällen 1 und 3 der Tabelle 5 zeigt, daß mittels der neun untersuchten Merkmale lediglich bezüglich einer euthyreoten gegenüber einer hyperthyreoten Symptomatik entschieden werden kann; die Ursache bleibt zunächst ungeklärt. (Tatsächlich wurden die beiden Patienten wegen starker hyperthyreoter Beschwerden untersucht.) Dekompensierte autonome Adenome, die sich ja hyperthyreot verhalten, werden durch das taxometrische Verfahren also als Hy eingestuft. Das taxometrische Verfahren erlaubt also, zwischen euthyreoten und hyperthyreoten Patienten zu unterscheiden; es klärt jedoch nicht die Pathogenese von Schilddrüsenstörungen, sondern unterscheidet lediglich zwischen zwei Schilddrüsenfunktionstypen; mit den verwendeten Merkmalen läßt

sich nur die hyper- bzw. euthyreote F u n k t i o n erkennen.

Nachdem Ausgangsstichprobe und Erprobungsstichprobe zusammengefaßt worden waren und so eine Verbesserung bei der Errechnung der Diagnosewahrscheinlichkeiten erzielt worden war, wurden auch die Wahrscheinlichkeiten für die drei Fälle von dekompensiertem autonomen Adenom neu berechnet (Tabelle 6).

Tabelle 6: Verbesserte Diagnosewahrscheinlichkeiten für die drei dekompensierten autonomen Adenome aus Tab. 5 (in %):

Fall	Hy	Ad	Eu
1	96	4	-
2	97	3	-
3	100	-	-

Die Wahrscheinlichkeit für Eu ist jetzt für alle drei Fälle praktisch gleich Null.

Diese Ergebnisse legen die Frage nahe, ob es nicht eigentlich um die Entscheidung zwischen nur zwei Hypothesen geht. Dekompensierte autonome Adenome (DAd) zeigen typisch hyperthyreote Symptome (Exophthalmus ausgenommen), so daß sich Hy und DAd zu der Diagnose Hy* zusammenfassen lassen. Andererseits lassen sich Euthyreose und kompensiertes autonomes Adenom (KAd) zu der Diagnose Eu* = Eu + KAd zusammenfassen. Dann wäre einfach zwischen Hy* und Eu* zu unterscheiden. Endgültig klären läßt sich diese Frage, falls bei größeren Stichproben die Ellipsen für Ad und Eu praktisch zusammenfallen.

Auch für die drei Fälle von dekompensiertem autonomem Adenom wurden die Diagnosewahrscheinlichkeiten bei Berücksichtigung nur der sieben klinischen Merkmale berechnet. Dabei ergab sich, daß nur Fall 3 mittels der klinischen Symptome als Hy^{*} zu entscheiden gewesen wäre (vgl. Tabelle 7).

Tabelle 7: Diagnosewahrscheinlichkeiten für drei dekompensierte autonome Adenome (in %) bei Berücksichtigung nur der sieben klinischen Merkmale:

Fall	Hy	Ad	Eu
1	28	24	48
2	51	24	25
3	100	-	-

Die Diagnose Hy^{*} wird also bei Berücksichtigung der Labordaten mit beträchtlich größerer Sicherheit getroffen.

3.3.4 Vergleich des Supermerkmals z₁ mit dem klinischen Eindruck

An dieser Stelle ist es interessant, auf die Ergebnisse der Untersuchung von IHM, KUNI, CHRISTIANSEN, JOSEPH und GRAUL (1970) näher einzugehen.

Die Schilddrüsenpatienten der nuklearmedizinischen Klinik und Poliklinik in Marburg/Lahn wurden einem Experten ohne Kenntnis der Untersuchungsbefunde vorgestellt, und der klinische Eindruck wurde nach folgender Skala objektiviert:

- 1 sicher euthyreot
- 2 deutlich vegetative Symptomatik mit euthyreoter Tendenz
- 3 stark vegetativ stigmatisiert
- 4 eher hyperthyreot als euthyreot
- 5 stark hyperthyreot
- 6 sehr stark hyperthyreot

Für jeden Patienten lag also ein Skalenwert und ein Wert z_1 für das aus den sieben klinischen Merkmalen ermittelte Supermerkmal vor. Skalenwerte und z_1 -Werte wurden miteinander korreliert; die Prüfung auf Linearität der Korrelation ergab Signifikanz ($p \leq 0.05$), d.h. Skalenwerte und z_1 -Werte stehen nicht in linearer Beziehung. Wurden jedoch nur die Fälle in die Linearitätsprüfung einbezogen, die aufgrund des klinischen Eindrucks den ersten vier Kategorien zugeordnet worden waren, so ließ sich Nichtlinearität nicht mehr nachweisen (vgl. Abbildung 15); z_1 stellt also ein Maß für Hyperthyreoidie dar.

Abbildung 15 zeigt, daß das Supermerkmal nicht linear dem klinischen Eindruck folgt, sondern auf Hyperthyreosen überproportional reagiert. Im Bereich der Euthyreosen ist diese Abweichung nicht zu beobachten. Daß hier die Steigung sehr flach verläuft, aus z_1 also der klinische Eindruck nur unsicher vorhergesagt werden kann, liegt sicher an der geringen Zahl von berücksichtigten Merkmalen.

Möglicherweise läßt sich die Nichtlinearität zwischen klinischem Eindruck und dem Supermerkmal z_1 mit dem WEBER'schen Gesetz erklären: "Beobachtet man Unterschiede zwischen zwei unterscheidbaren Größen,

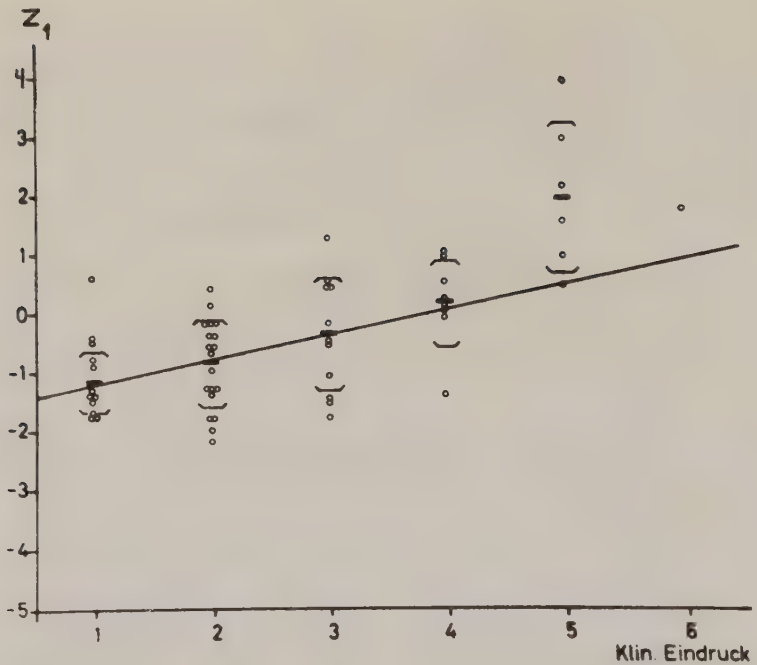


Abb. 15: Korrelation des Supermerkmals z_1 mit den Kategorien für den klinischen Eindruck. Angegeben sind Mittelwerte und einfache mittlere Streuung in jeder Kategorie mit Ausnahme von Kategorie 6, die nur einen Fall enthält und darum auch nicht in die Berechnung der Korrelation mit den Formeln für gruppierte Werte einbezogen wurde.

so erfassen wir das Verhältnis der beiden verglichenen Größen und nicht den Unterschied selbst." (DREVER/FRÖHLICH, 1970, 286) M.a.W.: bei heftigerer vegetativer Symptomatik bedarf es eines größeren Zuwachses an vegetativer Stigmatisierung, um die Eingruppierung in die nächst höhere Kategorie zu veranlassen. Bei Hyperthyreosen verfügt der Arzt demnach nicht über gute Differenzierungsmöglichkeiten für die Schwere der Erkrankung; obwohl die Kategorien 4 bis 6 dafür vorgesehen waren, finden sich bis auf zwei (je eine in 4 und 6) alle Hyperthyreosen in Kategorie 5.

Abgesehen von der überproportional hohen Differenzierungsfähigkeit des Supermerkmals z_1 im hyperthyreoten Bereich hat die Studie von IHM, KUNI, CHRISTIANSEN, JOSEPH und GRAUL also gezeigt, daß das Supermerkmal den Eindruck des Klinikers recht gut nachbildet; z_1 korreliert positiv mit der Skala, an der der erfahrene Arzt sich bei der Beurteilung von Schilddrüsenerkrankungen orientiert. Die Hauptkomponentenanalyse hat also den Faktor gefunden, den auch der Experte bei seiner diagnostischen Entscheidung empfindet; die Analyse des klinischen Eindrucks ergibt so eine Validierung des Hyperthyreoidie-Faktors.

3.3.5 Diagnosewahrscheinlichkeiten und Entscheidungsregeln

Es hat sich gezeigt, daß sich über die auf faktorenanalytischem Wege gefundenen Supermerkmale z_1 und z_2 mit Hilfe des BAYES'schen Theorems Diagnosewahrscheinlichkeiten berechnen lassen. Doch ist dem praktischen Arzt mit solchen Berechnungen nicht geholfen, wenn ihm keine Angaben über das Risiko einer Fehldiagnose

gemacht werden. Solche Risikobetrachtungen sind jedoch numerisch schwer anzustellen. Ein Ausweg aus dieser Schwierigkeit ergibt sich aus der Überlegung, wie der Arzt sich angesichts des Risikos von Fehldiagnosen verhält. Die Analyse des klinischen Eindrucks ergab, daß der Arzt sich an einer Skala orientiert, die dem gefundenen Supermerkmal z_1 recht gut entspricht. D.h. der Arzt fragt gar nicht nach Diagnosewahrscheinlichkeiten; er fragt vielmehr, bei welchen Patienten weiterführende nuklearmedizinische Diagnoseverfahren angebracht sind. Die Berechnung von Diagnosewahrscheinlichkeiten für sich löst noch nicht die Frage des praktischen Arztes, ob zusätzliche aufwendige diagnostische Untersuchungen wie Radiojodtest und Szintigramm an einem Patienten durchgeführt werden sollten.

Aus der Analyse des Arztverhaltens folgt, daß der Kliniker gar nicht zwischen zwei diskreten Hypothesen, also zwischen Euthyreose und Hyperthyreose, entscheidet; es geht ihm nicht um die Trennung von Diagnosen, sondern um die Trennung von ärztlichen Verhaltensweisen, nämlich darum, ob ein bestimmter Patient zu weiteren diagnostischen Untersuchungen in die Klinik einzubestellen ist oder nicht. Nicht die ärztliche Diagnose ist entscheidend, sondern die ärztliche Strategie.

Es geht also darum, zwei irrtümliche Entscheidungen zu vermeiden, nämlich 1. Eu fälschlich als Hy und 2. Hy fälschlich als Eu zu klassifizieren und entsprechend falsch zu behandeln. Beide Fälle würden einen Verlust bedeuten: im ersten Fall würde ein Patient unnötigerweise nuklearmedizinisch untersucht,

im zweiten Fall würde eine an sich angebrachte nuklearmedizinische Untersuchung unterlassen.

Demnach kommt es darauf an, auf dem Hyperthyreoidie-Index den Punkt zu finden, von dem an weitere nuklearmedizinische Untersuchungen bei minimalem Verlustrisiko anzustellen sind. Eine Möglichkeit besteht darin, das ärztliche Verhalten zu beobachten, d.h. die einzelnen Patienten ärztlich beurteilen zu lassen; gleichzeitig würde man die Diskriminanzfunktion berechnen. Als Grenze zwischen den Kategorien "sicher euthyreot", "deutliche vegetative Symptomatik..." usw. wird dann der Punkt angesehen, bei dem 50 % der Beurteiler sich für die untere und 50 % für die obere Kategorie entscheiden. Hierzu braucht man allerdings relativ große Fallzahlen.

Man kann aber auch den Regressionsansatz wählen, die Kategorien mit 1,2, ... 6 quantifizieren und für jeden Wert der Diskriminanzfunktion die Kategorie mittels der Methode der kleinsten Quadrate bestimmen (vgl. IHM, 1972).

In der Untersuchung von IHM, KUNI, CHRISTIANSEN, JOSEPH und GRAUL zeigte die Analyse des Arztverhaltens, daß der erfahrene Kliniker von der Kategorie 2 der Skala an ("deutlich vegetative Symptomatik mit euthyreoter Tendenz") alle Patienten zur Weiteruntersuchung mit dem Szintigramm einbestellt. Da das Supermerkmal z_1 mit der Urteilsskala des Klinikers korreliert, läßt sich die Entscheidung des Klinikers über lineare Regression nach Berechnung des z_1 -Wertes für einen Patienten unbekannter Diagnose vorhersagen. Es ist also möglich, Intervalle auf der z_1 -Achse zu definieren, die zu einzelnen Entscheidungen führen.

In Zukunft wird es darum gehen, nicht so sehr Wahrscheinlichkeitsangaben bezüglich bestimmter Diagnosen zu treffen, sondern zu Entscheidungsregeln zu kommen, die den Fortgang der diagnostischen Untersuchung oder der Therapie bestimmen. Solche Entscheidungsregeln beruhen zwar auf wahrscheinlichkeitstheoretischen Ansätzen, doch tritt die Wahrscheinlichkeit für den Benutzer nicht mehr unmittelbar in Erscheinung (vgl. IFM, 1972).

5. Diskussion

Zwei Verfahren zur automatischen Klassifizierung von Schilddrüsenerkrankungen wurden hinsichtlich ihrer Trennschärfe erprobt. Es ergaben sich folgende Schlüsse:

1. Das lernende Programm von ARKADJEW und BRAWERMAN ermöglicht keine befriedigenden Ergebnisse.

2. Die taxometrische Projektion erlaubt die Kontrolle von Ausreißern bzw. das Auffinden von Fehldiagnosen sowie die Entdeckung bisher unbekannter Krankheitseinheiten (PROKOPH, 1973).

3. Die Hauptkomponentenanalyse ergab einen Hyperthyreoidie-Faktor, der nicht die pathogenetischen, sondern lediglich die funktionalen Aspekte der Krankheitsbilder erfaßt. Dieser Faktor differenziert im oberen hyperthyreoten Bereich besser als der klinische Eindruck.

4. Die Berechnung von Diagnosewahrscheinlichkeiten aufgrund des Theorems von BAYES bleibt unbefriedigend; gefragt sind konkrete Entscheidungsregeln, die dem Arzt bei der Fortführung von Diagnose und Therapie helfen.

5. Die Analyse des ärztlichen Entscheidungsverhaltens ergab eine Korrelation zwischen der Urteilsskala des Klinikers und dem durch Hauptkomponentenanalyse gefundenen Hyperthyreoidie-Faktor. Damit ist es möglich, mit Hilfe des Supermerkmals z_1 die ärztliche Entscheidung vorherzusagen. Der Einsatz des Computers in der Schilddrüsendiagnostik ist also nicht für die Frage der Diagnose, sondern für die Frage der ärztlichen Strategie relevant.

5. Literaturverzeichnis

- ARKADJEW, A.G. und BRAWERMAN, E.M.: Zeichenerkennung und maschinelles Lernen, Beihefte zur Zeitschrift "Elektronische Rechenanlagen", Bd. 13, R. Oldenbourg Verlag, München und Wien 1966.
- DREVER, J. und FRÖHLICH, W.D.: Wörterbuch zur Psychologie, 3. Aufl. dtv, München 1970.
- GROSS, R.: Medizinische Diagnostik - Grundlagen und Praxis, Springer-Verlag Berlin, Heidelberg, New York 1969.
- HOFSTÄTTER, P.R.: Psychologie; Das Fischer Lexikon Bd. 6 Fischer Bücherei KG, Frankfurt am Main 1957.
- IHM, P.: Der Computer als Diagnostikhilfe dargestellt am Problem der Schilddrüsendiagnostik, Analele Universitatii Bucuresti Matematica - Mecanica, Anul XXI - Nr. 1 - 1972.
- IHM, P., TRAUTNER, R. und WOLF, H.: Linear algebraische Methoden in der numerischen Taxonomie, Biometrische Zeitschrift Bd. 13, 1971, H. 3, 161 - 202.
- IHM, P. und LIEBAU, A.: Homogenitätsprüfung viel-dimensionaler medizinischer Daten mittels Hauptachsentransformation, Methodik der Information in der Medizin, Bd. IV, Heft 3, 1965, S. 107 - 111.
- IHM, P., HIMMELMANN, G., HINZ, U., FÜRSCH, H.: Taxometrische Untersuchungen an Epilachna-Stichproben aus Zentralafrika, Biometrische Zeitschrift Bd. 9, Heft 3, 1967. S. 159 - 179.
- IHM, P., KUNI, H., CHRISTIANSEN, V., JOSEPH, K., GRAUL, J.: Elektronische Datenverarbeitung in der medizinischen Diagnostik am Beispiel der Hyperthyreose. Unveröffentlichtes Manuskript, Marburg/Lahn 1970.

- LÜLLWITZ, E.: Integration von Datenerfassung für EDV und Verwaltung in die ärztliche Kommunikation. Organisation der Arztbriefherstellung und Thesaurus für nuklearthyreologische Arztbriefe, Med. Diss. Marburg/Lahn 1972.
- McCOLLOUGH, und ATTA, : Programmierte Einführung in die Statistik.
- MITTENECKER, E.: Planung und statistische Auswertung von Experimenten, 4. Aufl., Wien 1963.
- PAWLIK, K.: Dimensionen des Verhaltens, Verlag Hans Huber, Bern und Stuttgart 1968.
- PRIES, H.H.: Integration von Datenerfassung für EDV und Verwaltung in die ärztliche Kommunikation. Organisation der Patientenaufnahme und Thesaurus für nuklearhämatologische Arztbriefe, med. Diss. Marburg/Lahn 1972.
- PROKOPH, J.: Programmierung des Rechenoperators TAXO 9 zur Taxometrie und seine Anwendung in der Schilddrüsendiagnostik, IMBIS (Inf. über med.-biol. Statistik) 2, Marburg 1973.
- SACHS, L.: Statistische Auswertungsmethoden, 2. Aufl. Springer Verlag, Berlin, Heidelberg, New York 1969.
- ÜBERLA, K.: Faktorenanalyse, Springer-Verlag, Berlin, Heidelberg, New York 1968.
- WEBER, E.: Grundriß der biologischen Statistik, VEB Gustav Fischer Verlag Jena, 6. Aufl. 1967.
- ZURMÜHL, R.: Matrizen. 2. Aufl., Springer-Verlag, Berlin, Göttingen, Heidelberg, 1958.

LEBENS LAUF

Als Tochter des Arztes Dr. med. Martin Schubert und seiner Ehefrau Dr. med. Margarete, geb. Hornig wurde ich, Veronika Schubert, am 4. Oktober 1943 als viertes von insgesamt vier Kindern in Cottbus/Niederlausitz geboren.

Im März 1945 mußten wir Cottbus verlassen; nach längerer Flucht gelangten wir nach Neuendorf/Schleswig Holstein. Dort wurde ich Ostern 1950 in die Volksschule eingeschult, wo ich die beiden ersten Volksschuljahre verbrachte. Danach ließ mein Vater sich als Facharzt für Chirurgie in Rellingen/Pinneberg nieder. Nach zwei weiteren Volksschuljahren in Rellingen bestand ich die Aufnahmeprüfung für das Lyzeum in Elms-horn, das ich bis zur Obersekunda besuchte; nach einem Wechsel zum Johannes-Brahms-Gymnasium in Pinneberg bestand ich im Februar 1964 das Abitur.

Mein Medizinstudium begann ich an der Christian-Albrecht-Universität in Kiel im Sommersemester 1964. Kiel verließ ich nach bestandenen Physikum im Wintersemester 1966. Während der klinischen Semester studierte ich in Marburg an der Philipps-Universität, wo ich im Sommersemester 1970 das Staatsexamen ablegte.

Im Dezember 1969 heiratete ich den Theologen und Psychologen Hauke Christiansen.

Meine Medizinalassistenzeit begann ich im Herbst 1970 im Allgemeinen Krankenhaus Wandsbek in Hamburg. Zum Oktober 1971 erhielt ich die Approbation als Ärztin. Anschließend arbeitete ich kurzfristig auf einer Anästhesieabteilung in Hamburg Barmbek, entschloß mich dann für die Facharztausbildung in Innerer Medizin, mit der ich im Januar 1972 im Allgemeinen Krankenhaus Altona in Hamburg begann.

Im Juli 1973 wurde uns eine Tochter, im August 1974 ein Sohn geboren.

1975 erhielt mein Mann eine Stelle in Bethel, was uns zum Umzug nach Bielefeld bewog. Hier schloße ich zur Zeit meine internistische Facharztausbildung an einem Kreiskrankenhaus ab.

Verzeichnis der akademischen Lehrer:

Prof.Dr.Andres, Prof.Dr.Bargmann, Prof.Dr.Bommer,
Prof.Dr.Buchholz, Prof.Dr.Berendes, Prof.Dr.Beckmann,
Prof.Dr.Cleve, Prof.Dr.Dölle, Prof.Dr.Exner,
Prof.Dr.Gedigk, Prof.Dr.Grewe, Prof.Dr.Hardewig,
Prof.Dr.Heiser, Prof.Dr.Herrlinger, Prof.Dr.Hildebrand,
Prof.Dr.Ihm, Prof.Dr.Jakob, Prof.Dr.Koransky, Prof
Prof.Dr.Kneeses, Prof.Dr.v.Kügelgen, Prof.Dr.Lang,
Prof.Dr.Leonhardt, Prof.Dr.Lullies, Prof.Dr.Lochte-Holtgreven,
Prof.Dr.Lindner, Prof.Dr.Linneweh, Prof.Dr.Martini,
Prof.Dr.Mann, Prof.Dr.Mannheim, Prof.Dr.Marx,
Prof.Dr.Neurath, Prof.Dr.Du-Mesnil, Prof.Dr.Ohlenbusch,
Prof.Dr.Overbeck, Prof.Dr.Pots, Prof.Dr.Puff,
Prof.Dr. Rohen, Prof.Dr.Schleyer, Prof.Dr. Schlosser,
Prof.Dr.Schirren, Prof.Dr.Schmidt, Prof.Dr.Schmitz-Moormann,
Prof.Dr.Schwaiger, Prof.Dr.Siegert, Prof.Dr.Straub,
Prof.Dr.Streicher, Prof.Dr.Strohmeyer, Prof.Dr.Voß,
Prof.Dr.Trinker.

Ich erkläre ehrenwörtlich, daß ich die am
Fachbereich Humanmedizin Marburg zur
Promotionsprüfung eingereichte Arbeit mit
dem Titel: "Entwicklung eines Verfahrens
zur automatischen Klassifizierung von
Schilddrüsenerkrankungen" im Institut für
Medizinische Statistik unter Leitung von
Prof. Dr. Ihm mit Unterstützung durch Prof.
Dr. Kuni, Prof. Dr. Graul, Dr. Joseph ohne
sonstige Hilfe selbst durchgeführt und bei
der Abfassung der Arbeit keine anderen als
die in der Dissertation angeführten Hilfs-
mittel benutzt habe.

Ich habe bisher an keiner in- und ausländischen
Medizinischen Fakultät ein Gesuch um Zulassung
zur Promotion eingereicht noch die vorliegende
oder eine andere Arbeit als Dissertation
vorgelegt.

Vorliegende Arbeit wurde nicht veröffentlicht.

Marburg, den 24. 11. 1978

Vorsnika Kunschhausen, geb. Schuber

